

Open Research Online

The Open University's repository of research publications and other research outputs

STELLAR (Semantic Technologies Enhancing the Lifecycle of Learning Resources): Jisc Final Report

Other

How to cite:

Brown, Sarah and Nurse, Richard (2013). STELLAR (Semantic Technologies Enhancing the Lifecycle of Learning Resources): Jisc Final Report. Jisc.

For guidance on citations see [FAQs](#).

© [not recorded]



<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Version: Version of Record

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data [policy](#) on reuse of materials please consult the policies page.

oro.open.ac.uk



Jisc Final Report

Project Information			
Project Identifier	<i>To be completed by Jisc</i>		
Project Title	STELLAR – Semantic Technologies Enhancing the Lifecycle of Learning Resources		
Project Hashtag	#OUstellar		
Start Date	1/02/2012	End Date	31/07/2013
Lead Institution	The Open University		
Project Director	Richard Nurse		
Project Manager	Sarah Brown		
Contact email	richard.nurse@open.ac.uk		
Partner Institutions	N/A		
Project Web URL	www.open.ac.uk/blogs/stellar		
Programme Name	<i>Digital Infrastructure</i>		
Programme Manager	Neil Grindley		

Document Information			
Author(s)	Sarah Brown, Richard Nurse		
Project Role(s)	Project Manager, Project Director		
Date	02/07/2013	Filename	Jisc STELLAR Final Report 2013-07. doc
URL	http://www.open.ac.uk/blogs/stellar/wp-content/uploads/2013/09/STELLAR-JISC-Final-Report.pdf		
Access	This report is for general dissemination		

Document History		
Version	Date	Comments
1.0	26/06/2013	First version. Input from Sarah Brown, Richard Nurse, James Alexander, Lara Whitelaw
2.0	02/07/2013	Version submitted to Jisc

Table of Contents

1	ACKNOWLEDGEMENTS	3
2	PROJECT SUMMARY	3
3	MAIN BODY OF REPORT	5
3.1	PROJECT OUTPUTS AND OUTCOMES	5
3.2	HOW DID YOU GO ABOUT ACHIEVING YOUR OUTPUTS / OUTCOMES?	7
3.2.1	Overview of approach	7
3.2.2	Workpackage 2 - Capturing perceptions	8
3.2.3	Workpackage 3 - Enhancement of the selected materials	11
3.2.4	Workpackage 4 - Evaluation	17
3.2.5	Workpackage 5 - Dissemination	18
3.3	WHAT DID YOU LEARN?	20
3.3.1	Project findings	20
3.3.2	Lessons learnt	27
3.4	IMMEDIATE IMPACT	30
3.5	FUTURE IMPACT	31
4	CONCLUSIONS	32
4.1	GENERAL	32
4.2	VALUE OF LEGACY LEARNING MATERIALS	32
4.3	APPLICATION OF SEMANTIC TECHNOLOGIES TO LEGACY LEARNING MATERIALS	33
4.4	SUSTAINABILITY OF USING SEMANTIC TECHNOLOGIES FOR SUCH COLLECTIONS	34
5	RECOMMENDATIONS	35
5.1	RECOMMENDATIONS FOR FUTURE DEVELOPMENT	35
5.2	RECOMMENDATIONS RELEVANT TO THE WIDER COMMUNITY	36
5.3	RECOMMENDATIONS FOR JISC	37
6	IMPLICATIONS FOR THE FUTURE	39
6.1	INTERNAL IMPLICATIONS	39
6.2	IMPLICATIONS FOR WIDER COMMUNITY	39
6.3	ONGOING AVAILABILITY OF PROJECT OUTPUTS	40
7	REFERENCES	41
8	APPENDICES	43
8.1	BALANCED SCORECARD	43
8.2	PRE-ENHANCEMENT SURVEY QUESTIONNAIRE	44
8.3	POST-ENHANCEMENT SURVEY	48
8.4	FEDORA XCRI MODULE RECORD FOR S100	50
8.5	FEDORA SAMPLE RECORD FOR MAIN TEXT	56
8.6	MODULE PROFILE	59
8.7	OU STUDY MATERIAL OBJECT MODEL	66
8.8	DRAFT OPEN UNIVERSITY DIGITAL PRESERVATION POLICY (AS OF 26/06/2013)	67
8.9	DRAFT OPEN UNIVERSITY DIGITAL PRESERVATION STRATEGY AND SELECTION CRITERIA (AS OF 26/06/2013)	71

1 Acknowledgements

1. The STELLAR (Semantic Technology Enhancing the Lifecycle of Learning Resources) project was funded by Jisc as part of the Digital Infrastructure programme, in the 'Enhancing the Sustainability of Digital Collections' strand of the 16/11 Jisc Grant funding stream¹. The project ran for eighteen months and completed on 31 July 2013.
2. The project is grateful to the Jisc for the opportunity to carry out this piece of work on a unique collection of digital materials and for the support of the Programme Manager Neil Grindley throughout the project. Whilst the STELLAR project was run from within the Open University Library Services and most of the project team was drawn from that Unit we were pleased to have the involvement of several colleagues from across The Open University, both on the Project team and the Project Steering Group. The project would particularly like to thank Dr Mathieu d'Aquin, Research Fellow, at the Open University Knowledge Media Institute² for his work to create the semantic web developments for the project and for permission to build on and adapt his development of the DiscOU semantic web tool³ for the STELLAR project.
3. The project would also like to thank members of the Project Steering Group, particularly Liz Burton-Pye (Head of Learning and Teaching), Dr Shafquat Towheed (Lecturer in English) and Stuart Brown (eBusiness Manager – Online Communities & Web).

2 Project Summary

4. As one of the earliest distance learning providers The Open University (OU) has a rich heritage of archived learning materials. An ever increasing amount of that is in digital form and is being deposited with the University Archive. This growth has been driven by digitisation activity from projects such as AVA (Access to Video Assets)⁴ and the Fedora-based⁵ Open University Digital Library⁶ '*a place to discover digital and digitised archival content from the OU Library, from videos and images to digitised documents*'⁷. Other digital content is being captured from web archiving activities, such as work to preserve Moodle⁸ Virtual Learning Environment course websites. An evidence based understanding is required

¹ http://www.jisc.ac.uk/fundingopportunities/funding_calls/2011/10/grantcall1611.aspx

² <http://kmi.open.ac.uk/>

³ <http://kmi.open.ac.uk/technologies/name/discou>

⁴ <http://www.open.ac.uk/blogs/AVA/>

⁵ <http://fedora-commons.org/>

⁶ www.open.ac.uk/blogs/oudl

⁷ http://www.open.ac.uk/blogs/OU DL/?page_id=141

⁸ <https://moodle.org/>

to inform digital preservation policies, curation strategy and investment in digital library development.

5. Following the Pre-enhancement, Enhancement and Post-enhancement methodology set out by Jisc, STELLAR adopted the model of a balanced scorecard to ascertain the value ascribed to the non-current learning materials. Four aspects were considered: Personal and professional perspectives of value; Value to the Higher Educational and academic communities; Value to internal processes and cultures; Financial perspectives of value. The outcomes of the survey⁹ indicated that stakeholders place a high value on the materials, and that they perceived them to have value in all areas evaluated.
6. Three OU courses were chosen from the digital library for the transformation stage. These materials were enhanced and transformed into RDF, a process that required more extensive metadata expertise and effort than was expected. Following enhancement the RDF was accessed through a tool called DiscOU, created by a member of the project team from the OU's Knowledge Media Institute. DiscOU uses both linked data and a semantic meaning engine to analyse the meaning of the text in a search query. This is matched against the meaning of the content derived from an index of the full-text of the digital library content.
7. In the final stage stakeholders were asked through a survey and series of workshops to use the DiscOU proof-of-concept tool to assess their perception of the value of this transformation. This has revealed that overall, academics and other stakeholders in the university do believe that the value of the selected materials was positively impacted by the application of semantic technologies.

⁹ <http://www.open.ac.uk/blogs/stellar/wp-content/uploads/2012/10/STELLAR-Project-Pre-enhancement-survey-report.pdf>

3 Main Body of Report

3.1 Project Outputs and Outcomes

Output / Outcome Type (e.g. report, publication, software, knowledge built)	Brief Description and URLs (where applicable)
Report	<p>A pre-enhancement benchmark survey report detailing analysis of results and informing the selection of material to be chosen for conversion into linked data format</p> <p>http://www.open.ac.uk/blogs/stellar/wp-content/uploads/2012/10/STELLAR-Project-Pre-enhancement-survey-report.pdf</p>
Software	<p>A proof of concept to demonstrate the enhancement of a sample of learning materials using linked data within the Open University Digital Library</p> <p><i>A proof-of-concept tool has been created based on the DiscOU engine and has been used to demonstrate the enhancement.</i></p> <p>http://www.open.ac.uk/blogs/stellar/?p=158</p> <p><i>Linked data capabilities have been embedded into the OU Digital Library making use of Fedora's linked data capabilities. SPARQL queries have also been utilised within the OU Digital Library to surface relevant content.</i></p>
Report	<p>A post-enhancement activity survey report</p> <p>http://www.open.ac.uk/blogs/stellar/wp-content/uploads/2013/07/STELLAR-Post-Enhancement-Survey-Report.pdf</p>
Report	<p>Report on the sustainability of using linked data for legacy learning materials.</p> <p><i>Covered in Section 4.4 of the Jisc Final report</i></p>
Knowledge built	<p>Documentation of an evidence-based model for lifecycle management, including recommendations for a) Digital preservation policy and b) Digital Library development</p> <p><i>A draft OU Digital Preservation Policy and selection criteria for content for the Digital Library are included as appendices to this Jisc Final report in sections 8.8 and 8.9</i></p>
Publication	<p>Project blog (www.open.ac.uk/blogs/stellar) and project twitter tag #oustellar</p> <p>www.open.ac.uk/blogs/stellar We aim to maintain this blog for at least five years and intend to web archive the content within the Digital Library</p>

Publication	<p>Internal dissemination activities</p> <p><i>A wide range of activities</i></p> <ul style="list-style-type: none"> • <i>Library Services staff development hours</i> • <i>Library Services Show & Tell display</i> • <i>Articles on Library Services intranet</i> • <i>Postings on Library Chat (Library Services blog)</i> • <i>Articles on OU intranet</i> • <i>Meetings with stakeholders across the university</i> • <i>Science Open Forum presentation</i> • <i>Evaluation workshops</i> • <i>Article in e-Learning Digest newsletter</i> • <i>Periodic email updates to individuals who expressed an interest in the project</i> • <i>Report to Learning and Teaching Strategy Group</i> • <i>Reports and presentations to Pro-Vice Chancellors</i>
Publication	<p>Participation in Jisc Programme events Presentations and papers at appropriate dissemination opportunities, Promotion via library, OU and personal channels:</p> <ul style="list-style-type: none"> • <i>Jisc programme meetings attended throughout project</i> • <i>Article in SCURL newsletter</i> • <i>Inclusion in SCONUL Focus 55</i> • <i>Presentation at ELAG 2013</i>
Knowledge built	<p>A clearer understanding of the value proposition for legacy learning materials at the OU that may have relevance and traction with the wider HE community</p> <p><i>The pre-enhancement report gives a clear articulation of the value of the materials and goes far beyond the previous perception that value was largely around re-use potential. We have reported the findings of the report to key stakeholders within the Institution and are using it as a key piece of evidence in our planning and policymaking.</i></p>
Knowledge built	<p>A clear set of selection criteria to inform the selection of legacy learning material for preservation. This will add to the body of evidence for the HE sector.</p> <p><i>Selection criteria included within section 8.9 as part of the Digital Preservation Strategy draft document.</i></p>
Publication	<p>Sample collections of legacy learning materials in linked data format (and connected to other linked data sources)</p> <p><i>Sample collections from OU courses: S100, D251, L550 transformed into linked data.</i></p>
Knowledge built	<p>An increased understanding of the processes needed to create and maintain linked data and of the value of linked data for transforming and accessing legacy learning materials.</p>

	<p><i>The project has been invaluable in allowing us to explore this area. We have created metadata ontologies that are being utilised within our Digital Library and will have linked data production embedded in our Digital Library processes. We understand clearly the effort needed to establish profiles for new types of material and have used that knowledge in our digital library planning.</i></p> <p><i>We have also built significant understanding of where the value lies in using linked data to access legacy learning materials. We have also been able to start to open up conversations about how this type of development could be applied during the course production Stage Gate process to help course creators find material for re-use from wider sources of content.</i></p>
Report	<p>Some of the sustainability aspects of linked data for this type of material have started to be exposed.</p> <p><i>The sustainability section of this report (4.4) and the recommendations discusses this aspect. Extensive metadata effort is required at an initial stage, and to 'clean-up' data. High-quality metadata is needed for linked data. Once this stage of work is completed it is possible to structure systems to create linked data routinely with little on-going effort as long as metadata standards are maintained.</i></p> <p><i>Technical transformation activities are becoming easier as the transformation becomes more routine and tools are developed. Additional work within Fedora to support linked data developments would be welcomed.</i></p>

3.2 How did you go about achieving your outputs / outcomes?

3.2.1 Overview of approach

8. The overall objectives of the STELLAR project were to:
 - a. Develop a detailed understanding of the value of legacy learning materials as perceived by academic staff and other key stakeholders
 - b. Using the evidence gained in the surveying phase experiment with the use of semantic technologies in a digital library environment to ascertain the extent to which the perceived value of these materials might be enhanced and to consider the sustainability implications of using semantic technologies.
 - c. Inform the development of digital libraries of learning resources by contributing to the evidence base for their effectiveness
 - d. Increase the return on investment of learning materials by developing an evidence based model for lifecycle management

- e. Disseminate the experience of the project through workshops, conference papers, journal articles and blogs
9. To achieve these objectives the project was organised into three main areas of activity:
- A benchmarking survey of academic staff, plus interviews with selected stakeholders to investigate the value they placed on legacy learning materials
 - Informed by the survey the project selected material for the experimental application of semantic technology/linked data approaches to try and enhance the value as expressed in terms of the discoverability, flexibility, and the ability to disaggregate and reuse the content
 - An evaluation to ascertain the effectiveness of these approaches.
10. The project activities were split into five work packages. Work packages 1 and 5 dealt with the Project Management, dissemination and engagement activities. Project work package 2 comprised the pre-enhancement surveying activity which identified the value that academics place on legacy learning materials. This then informed work package 3, which focused on the transformation of the material into linked data. Work package 4 looked at how that transformation affected the value of that material.

3.2.2 Workpackage 2 - Capturing perceptions

11. Whilst there is an increasing interest in the reuse of legacy learning materials there has been no systematic research into understanding their value. Studies such as eSPIDA¹⁰ and the Blue Ribbon Task Force report¹¹ have largely focused on the high level economic and strategic perspectives. Consultancy work carried out in 2008 by Kings Digital Consultancy Services to investigate the feasibility of digitising and preserving the OU's learning materials archive (Tanner, KCL, 2008 *Open University Footsteps Project – Investigating the Feasibility of digitising and processing the Open University Course Materials Archive*), emphasised financial value rather than the academic perspective.
12. This first phase of the project sought to ascertain the extent to which academic staff and stakeholders at the OU value these materials, whether that value is based on pedagogic or other reasons, and what they saw as the barriers to realising that value.

¹⁰ <http://www.gla.ac.uk/espida/index.shtml>

¹¹ http://brtf.sdsc.edu/biblio/BRTF_Final_Report.pdf

13. The first important decision made by the team was the definition of value – a broad definition was adopted. As Peter McKinney, of the eSPIDA¹² project at Glasgow University, has stated: “The key lies in value and an ability to articulate all areas of value that digital objects have. This means grappling with the concept of intangible value and expressing it in a way that is as powerful as purely monetary arguments. The business world has been aware that the value and worth of companies has long since shifted from the accounting of only tangible assets to the inclusion of intangible assets such as information capital, processes and human capital” (McKinney, 2005: pp.1-2)¹³. The STELLAR Project adapted the balanced scorecard approach¹⁴, recommended in the findings of the Jisc-funded eSPIDA project, to the context of the Open University, in order to ensure that a full picture of value was established.
14. The four perspectives of value that this approach allowed STELLAR to tap into were:
- Personal and professional perspectives of value
 - Value to the Higher Educational and academic communities
 - Value to internal processes and cultures
 - Financial perspectives of value.
15. The second important decision was that of whose opinions to canvas. In order to gain as complete a picture as possible, and to ensure that the findings had the greatest potential impact within, and beyond, the confines of the STELLAR project, three groups were selected for survey. Firstly, asset creators – those who are directly involved in the production of learning materials at the OU. This includes academic staff and non-academic staff, such as Curriculum Managers, who are involved in a higher administrative capacity. Secondly, senior stakeholders – those with senior roles within the university who may have a specific departmental or faculty interest in learning materials, or who may be involved in drafting future policy or in fund allocation to projects concerning their preservation. Thirdly, those who use learning materials on a regular basis. Surveying students was beyond the scope and resource of this study, as legacy course material collections are not open to student access, but surveying Associate Lecturers¹⁵ (ALs) was not. It is ALs that make-up the end-user element of the population under study.

¹² www.gla.ac.uk/espida

¹³ McKinney, P., 2005. *eSPIDA and Sustainable Digital Preservation*. Society of Archivists: <http://hdl.handle.net/1905/443>

¹⁴ http://en.wikipedia.org/wiki/Balanced_scorecard

¹⁵ Associate Lecturer is the term used for the majority of staff who teach Open University students. They support students by making contact at key points in their study, marking and providing feedback on assignments, helping students understand the module material and helping them prepare for examinations or end of module assessments.

16. Data was collected from this population in two ways. Firstly through an online survey, made available to the entire population and, secondly, through a number of follow up interviews carried out with a sample of those who had filled in the survey.
17. The survey was designed to be easy to fill in, in order to maximise response rate and, as such, a Likert-type scale¹⁶ was adopted, with only one opportunity, at the end of the survey, for participants to leave a written response. The number of Likert items was kept to a minimum and they were designed to tap into each of the four perspectives of value outlined above. The online survey was advertised throughout the university in, for example, online newsletters; course meetings; departmental and faculty emails; the university intranet; and through the STELLAR team twitter accounts and blog.
18. Asset creators, senior stakeholders and regular users of such material at the OU were invited to respond to the survey using a number of distribution channels. There were 561 responses to the survey.
19. Follow up interviews were designed to be brief, lasting between 15 and 20 minutes, and were recorded for post-hoc analysis. An interview schedule was designed to keep the conversations focussed, but to enable a more free ranging exploration of the four perspectives of value than the surveys allowed, with the aim of adding nuance to the data already collected. Stratified sampling was used to select the 34 interviewees. The population numbered 299 - those who had indicated that they would be willing to take part in a follow up interview in their survey response – and the strata, the different departments, units and faculties that the respondent had indicated were their primary place of work within the OU. Approximately 10% of each strata were randomly selected for interview, and all agreed. 23 of the interviews were conducted face-to-face, while the remaining 11 were conducted over the phone. Twelve senior stakeholders, who had not completed the online survey, were identified and invited to take part. Six of these agreed and were interviewed in addition to the 34 follow-up interviews.
20. Data from both the surveys and interviews were anonymised during the analysis and identifiable data was only kept where necessary – for example where an expression of interest in continuing to be involved with the STELLAR Project, was made.

¹⁶ http://en.wikipedia.org/wiki/Likert_scale

21. The full report for the pre-enhancement survey has been published on the Stellar blog¹⁷. The report has been used not only within the project but also to inform wider OU policy discussions about the value of non-current learning materials. The investigation of each of these dimensions served as a benchmark against which to test the effects of the transformation to be carried out in phase two of the project.

3.2.3 Workpackage 3 - Enhancement of the selected materials

3.2.3.1 Selection

22. Based on the findings from the survey, three collections were selected for enhancement in the second phase. They were selected to cover a broad range of aspects; historical significance, academic content and the variety of format types included. The chosen non-current OU modules were:
- S100 Science Foundation Course – presented from 1971 to 1978
 - D251 Issues in Deafness – presented from 1991 to 1996
 - L550 World Class Business English (an OU pack, or non-taught module) – available from 1999 to 2010
23. For S100 main text print and AV study materials were already available in digital form as part of the digital library collection. The STELLAR team discussed whether it would be valuable to include other material associated with S100, such as Assessments. This would provide greater depth to the collection and would allow the project to test the viability of handling a complete non-current OU module and transforming that into Linked Data. It was decided that this would be a useful exercise and the Digital Library team was asked to undertake some further digitisation to include all print supplementary items in the Digital Library; including content such as Tutor-marked Assignments (TMAs), Computer-marked Assignments (CMAs), Self Assessments and Summer School Notes. As these supplementary items were not previously catalogued, new catalogue records were created for these items. In addition to this task, in some cases improvements were required to existing catalogue data to support the transformation to Linked Data, adding information in relation to the collection the resources belonged to, its learning resource type classification, and a consistent location for the module code.

¹⁷ <http://www.open.ac.uk/blogs/stellar/wp-content/uploads/2012/10/STELLAR-Project-Pre-enhancement-survey-report.pdf>

3.2.3.2 Metadata

24. Based on the approach taken to the module materials it was decided that it was necessary to be able to view the metadata for the module resources from a module-based perspective; to this end it was decided that a module profile should be developed.

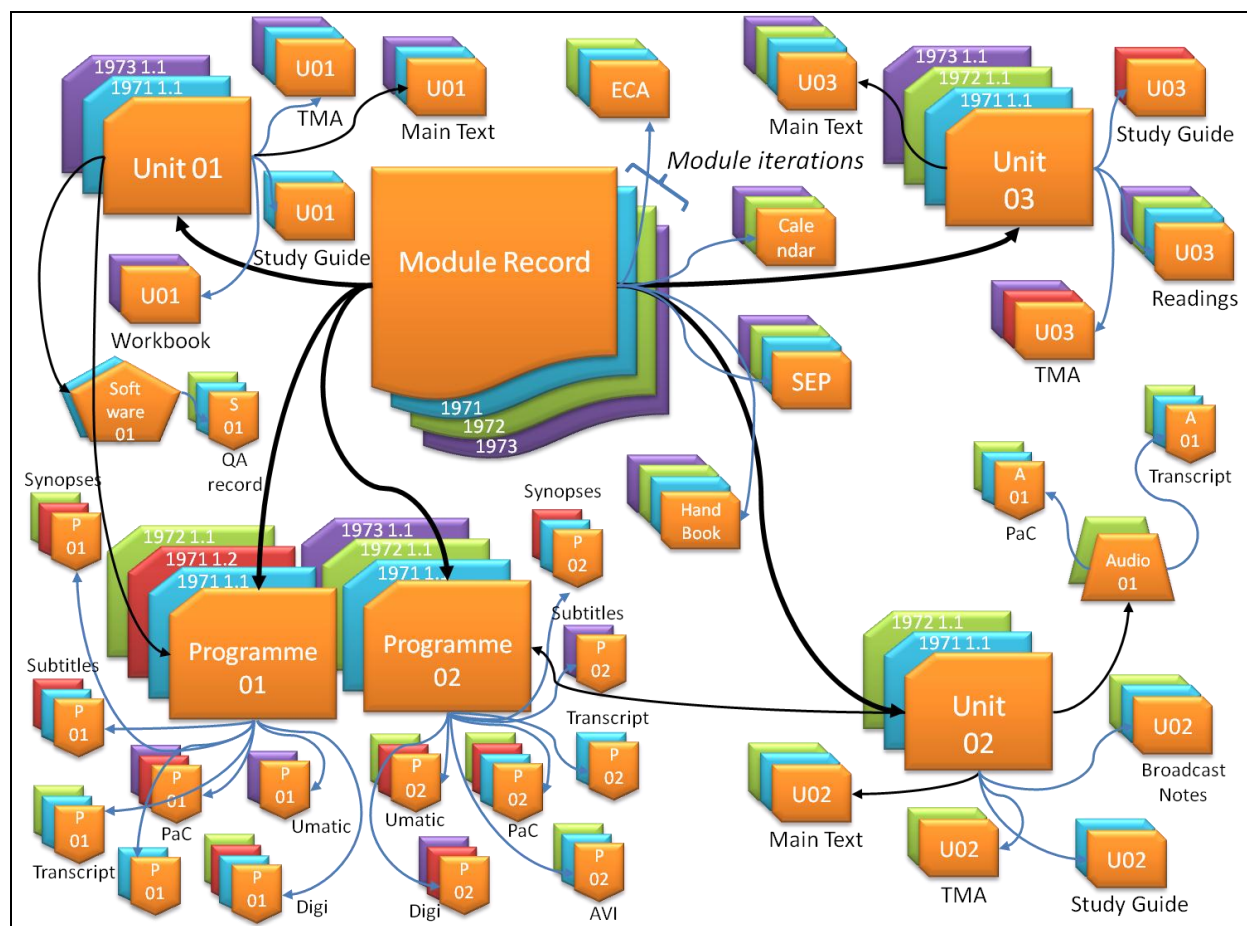


Figure 1: Diagram representing the module profile

25. Focusing on the need to express this profile as Linked Data, several different methods of expressing module/course metadata, were assessed, including Courseware¹⁸ and MLO¹⁹. It was decided to use the eXchanging Course Related Information (XCRI)²⁰ profile for this data type (see 8.4 Fedora XCRI module record for S100). XCRI was the most comprehensive of the schemas, and is promoted in Jisc and is being used for the HEFCE funded Course Data

¹⁸ <http://courseware.rkbexplorer.com/>

¹⁹ <http://www.cen-itsa.net/Main.aspx?put=1042>

²⁰ <http://www.xcri.co.uk/>

Programme²¹. Data the university records in its curriculum management system was utilised, along with library catalogue data and the module-related data the university creates for prospectus and module advertising, to populate these records.

26. The metadata pertaining to module resources is housed in the university library's (MARC-based) Voyager²² catalogue, or already stored within the OU Digital Library (OUDL). To make these records more usable within a module-based educational framework which needed to be expressible as RDF, the records for all text-based assets were mapped and transformed into Metadata Object Description Schema (MODS) and utilised the MODS RDF ontology²³. Audio-visual assets were mapped to the European Broadcast Union (EBU) EBUCore²⁴ metadata profile and used the EBUCore ontology²⁵.
27. A relationship model was then created (see 8.7 OU study materials object model) to express the connections within this profile. This enabled the resources to be presented in ways that were more logical and semantically sound than the standard catalogue interfaces previously available.

3.2.3.3 OUDL development

28. For the data enhancement, where records already existed detailing relationships between content items, regular expressions were used to re-write this as RDF into Fedora RELS-EXT datastreams. The process was facilitated by scripts for bulk ingest and purge of records and by Fedora's in-built capability for turning Dublin Core and RELS-EXT datastream relationships expressed in RDF into triples in its internal Mulgara triplestore on ingest.
29. However, where records did not exist or did not contain the necessary relationships, the process required looping through sets of records, recording the related identifiers and writing them to file (e.g. pages of supplementary items have to be related to their 'parent record', such as an exam paper). Up to five complete parses of data were necessary together with some matching by hand. When the 'round-tripped' record relationships were created (ie 'Y hasPart X' corollary 'X is a part of Y') SPARQL queries were used on the enhanced data.
30. Existing records of courses from a relational database were extracted by a series of queries and then transformed into records modelled in XCRI.

²¹ <http://www.xcri.co.uk/the-xcri-capital-programme.html>

²² <http://www.exlibrisgroup.com/category/Voyager>

²³ <http://www.loc.gov/standards/mods/modsrdf/primer.html>

²⁴ http://tech.ebu.ch/docs/tech/tech3293v1_4.pdf

²⁵ <http://www.ebu.ch/metadata/ontologies/ebucore/>

3.2.3.4 Proof-of-concept STELLAR tool

31. In order to demonstrate the linked data work, a semantic web tool was created to search the linked data representations of the digital library. To create the proof-of-concept tool STELLAR used an application called DiscOU²⁶ developed in the OU Knowledge Media Institute (KMi) which builds on previous work including the linked data materials created for the Jisc-funded Lucero²⁷ project that established data.open.ac.uk.
32. The interaction with the tool is similar to the one of DiscOU alfa²⁸. Starting from some text (preferably a couple of sentences or a paragraph), the tool will analyse the content of the text to extract the key, relevant concepts (entities such as people, places, topics, etc.) in this text. It then tries to find in the indexed course material, resources that cover those concepts to a similar degree. It then shows the concept extracted, ranked according to their importance in the text, and the list of found resources, with the metadata being displayed from their linked data-based descriptions. For each resource, it also shows the concept which most prominently connects the text and the resources, and the weight of concepts can be customised, to instruct the tool to give more or less importance to them according to the user's own interest. The display of the found resources links directly to the corresponding PDFs (e.g. scanned pages of study material) in the OUDL. Information about the course is obtained from and linked to data.open.ac.uk.

²⁶ <http://discou.info/>

²⁷ <http://lucero-project.info/>

²⁸ <http://discou.info/alfa/>

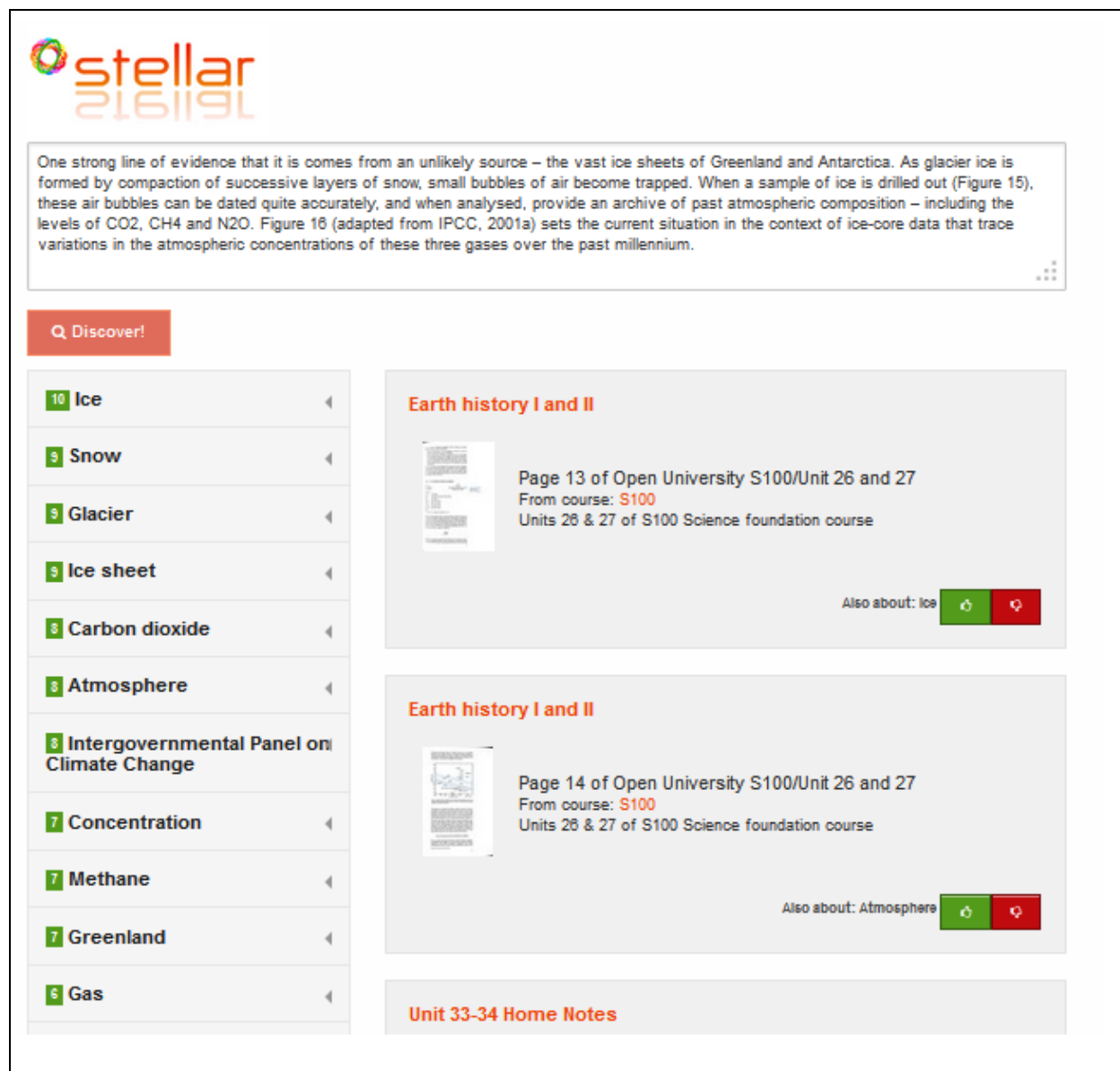


Figure 2: Screenshot of STELLAR proof-of-concept tool

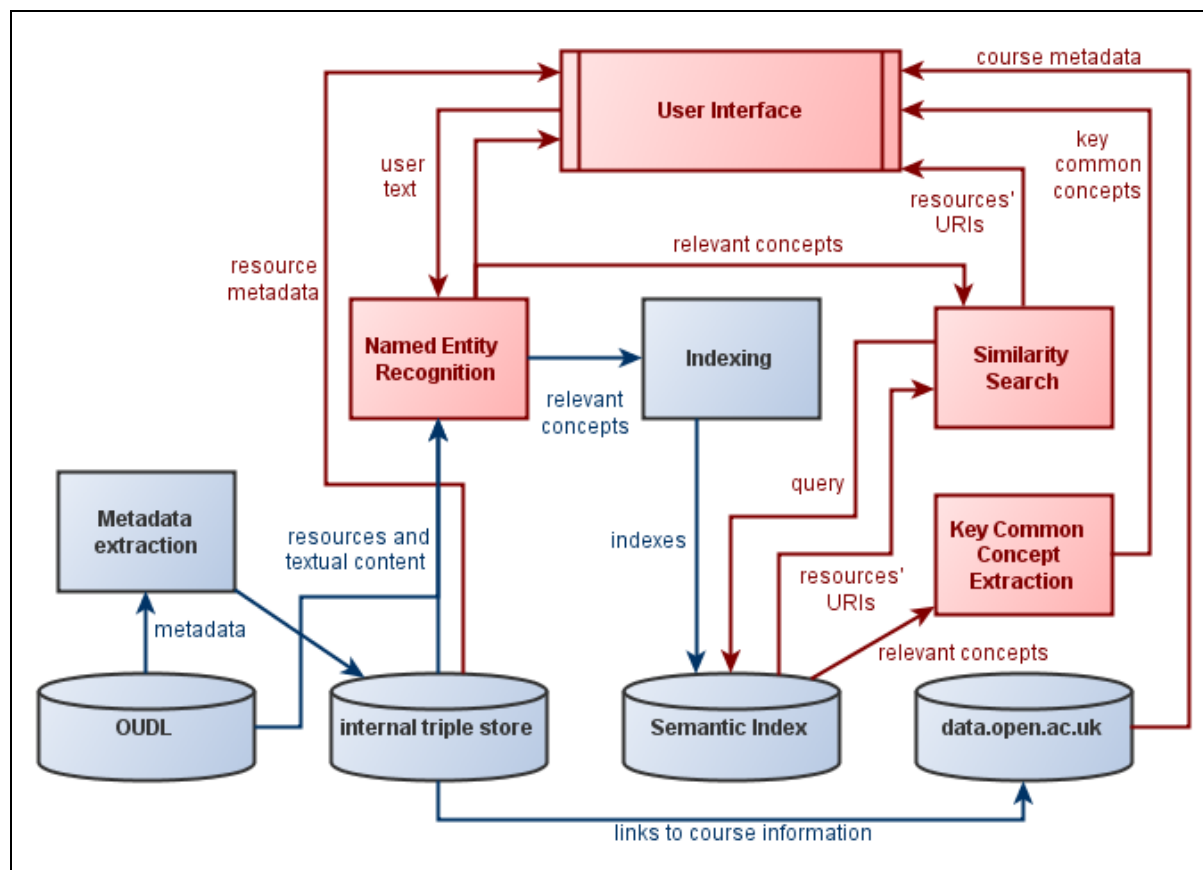
33. The backend of the tool is the generic DiscOU semantic similarity search engine, specially adapted with an index of the resources considered in STELLAR. Metadata for and links to each resource (page of study material, etc.) are first extracted from the OUDL, and transferred into a dedicated “intranet” triple-store connecting to data.open.ac.uk. These resources are then crawled, and their textual content extracted. A named entity recognition system (DBpedia Spotlight²⁹) is used to extract the relevant concepts from this text as linked data entities (in DBpedia³⁰). An index is then created using the Lucene open source search

²⁹ <http://spotlight.dbpedia.org/>

³⁰ <http://dbpedia.org/>

library³¹ that associates each resource with the relevant weighted entities. The information infrastructure of the STELLAR tool comprises this index, the “intranet” triple-store with dedicated metadata and data.open.ac.uk. It is worth mentioning as well that, as a side effect of STELLAR’s work on the module metadata, information about courses no longer in presentation at The Open University has been added to data.open.ac.uk.

34. The search mechanism and the instantiation of the interface of the STELLAR tool are achieved through a series of dedicated services. The text given by the user is first analysed using DBpedia Spotlight as in the indexing process, and the relevance scores normalised to provide the set of key relevant concepts. These are shown on the left hand side of the interface. The weighted concepts (possibly customised by the user) are then sent to another service which uses the search by similarity function of Lucene on the previously created semantic index (the “MoreLikeThis” search mechanism). Once the list of (URIs of) similar resources has been obtained from Lucene, the two linked data stores are queried to obtain all the relevant metadata. Another service is also queried to compute from the set of concepts shared between the query text and the return resources the one that should be considered most important.



³¹ <http://lucene.apache.org/>

Figure 3: Overview of the architecture of the STELLAR tool

3.2.4 Workpackage 4 - Evaluation

35. Once the materials were enhanced it was necessary to evaluate whether the transformation had increased the perceived value of the legacy materials. This evaluation was focused particularly on those individuals who had responded to the first survey and indicated they would like to remain involved with the project. This was to ensure that participants were already familiar with the objectives of the STELLAR project and up-to-date with progress through the ongoing dissemination activities.
36. The evaluation was conducted in three main ways:
37. Face to face workshops with small groups.
Participants in the workshops were presented with an overview of the project approach and progress to date to refresh their knowledge and understanding. The STELLAR proof-concept-tool was then shown to them to demonstrate how the selected course materials had been enhanced with linked data. They were then presented with the results of the semantic analysis carried out by the tool. Following this they were given a set of examples and invited to try the tool for themselves, entering the suggested text into the tool and viewing the content recommended in the results. If participants wished to enter their own examples they were free to do this on the understanding that only a selected set of materials had been enhanced and their text should reflect the themes of the three courses to provide meaningful results.
38. Demonstration to the Science Faculty.
The demonstration was delivered through a Science Open Forum and took a similar format to the workshops, but did not have the hands-on element as it was delivered in a presentation style. However attendees were encouraged to ask questions and there was a good level of discussion. Science colleagues engaged particularly well with the demonstration as much of the content available via the STELLAR tool came from the S100 Science Foundation course. This made it easier for this audience to consider the value and reusability of the material suggested by the tool as it was highly relevant to them
39. Opportunity to review the materials using the STELLAR proof-of-concept tool online.
To enable individuals to take part in the evaluation from their desk it was made available to OU staff via the project blog³². Participants were encouraged to read a blog posting describing recent project developments and containing an overview of the STELLAR proof-of-

³² http://www.open.ac.uk/blogs/stellar/?page_id=153

concept tool. They were then invited to experience the tool for themselves, using a set of examples provided to them. Again these examples were intended to return meaningful results in order that participants could see the impact of the linked data and how the semantic analysis of the text entered had worked. An invitation to take part in the online evaluation was sent via email to all individuals who had participated in the pre-enhancement survey and indicated they wished to remain involved in the project. It was also posted on the OU intranet and publicised in the e-Learning Digest newsletter.

40. As part of each of these events, participants were invited to complete the same questionnaire (see section 8.3) to determine whether the application of semantic technologies has enhanced the value of the materials in any way. As in the original pre-enhancement survey, the questions contained in the evaluation were designed to link to the 4 perspectives of value represented on the balanced scorecard.

3.2.5 Workpackage 5 - Dissemination

41. As part of the project planning, the key stakeholders in the STELLAR project were identified. The dissemination plan was then developed to meet the communication needs of these particular individuals and groups. Stakeholders were both internal to the OU (e.g. academics, academic-related staff, Associate Lecturers, Library Services staff, faculties, departments) and external (e.g. Jisc, other Library and Library Systems staff, the HE community)
42. Communication and dissemination activities continued throughout the project and were intended to:
- Raise awareness of the STELLAR project both inside and outside the OU
 - Engage stakeholders
 - Share project progress and findings
43. A key vehicle in dissemination of project information to all stakeholders was the project blog³³. This was updated regularly with information on progress and when key milestones in the project plan were achieved. Major deliverables such as the pre-enhancement survey were shared via the blog as well as being circulated directly to key stakeholders.

3.2.5.1 Internal dissemination

³³ <http://www.open.ac.uk/blogs/stellar/>

44. Internal to the OU a number of communication activities were undertaken. Presentations were made to Library Services staff to inform them about the project and updates were made to OU faculties, units and other colleagues engaged in similar projects. Activities such as the pre-enhancement benchmarking survey and the final evaluation workshops and demonstrations also provided an opportunity to communicate about the project and its findings. Information about the project was posted on the university intranet at appropriate intervals and included in the e-Learning Digest (internal newsletter).
45. Those individuals who had expressed an ongoing interest in the project were kept-up-to date via email and directed to updates on the project blog. We also disseminated key documents in the form of summary reports and presentations to important strategic groups such as the Learning and Teaching Strategy Group and to key stakeholders such as Pro-Vice Chancellors.

3.2.5.2 External dissemination

46. The project blog was the main vehicle for sharing information about the project externally. However, other specific activities were carried out.
47. For example, the findings of the pre-enhancement survey were shared with SCURL³⁴ in order to communicate to other organisations the value placed by OU academics on our legacy learning materials. An item on STELLAR was also included in SCONUL Focus 55³⁵. An overview of the project and its findings was presented by the Project Manager at the ELAG 2013³⁶ conference in May 2013. The presentation was delivered to the around 200 conference attendees and a similar number who viewed online³⁷. The video³⁸ and slides³⁹ are available for this presentation. We plan to continue dissemination activities in the next year to promote the findings of the project and take forward next steps.
48. The project team attended Jisc programme meetings to share project progress with and gather feedback from other institutions.

³⁴ <http://scurl.ac.uk/>

³⁵ <http://www.sconul.ac.uk/sites/default/files/documents/17.pdf>

³⁶ European Library Automation Group <http://www.elag.org/homepage.htm>

³⁷ <http://www.open.ac.uk/blogs/stellar/?p=246>

³⁸ http://youtu.be/ChET_-pCkyY

³⁹ <http://elag2013.files.wordpress.com/2013/06/elag-conference-paper-may-2013-v0-3.zip>

3.3 What did you learn?

3.3.1 Project findings

3.3.1.1 General findings

49. The project consumed more metadata resource than expected (as outlined in 3.3.1.5.2 below). This was both at the level of establishing the metadata schemas and also at the level of needing to carry out additional metadata cataloguing activity. The project team spent a total of 45 days mapping the metadata schema, which was significantly more than the 6 days originally estimated in the project plan. In addition, the Digital Library team undertook 11 days of work as part of their work programme to ensure that material needed for the STELLAR project was enhanced by completing the extra cataloguing required for the supplementary items.
50. This additional effort has to be considered as a one-off activity and the extent of it was partly driven by complexity but also due to the scope and range of the different types of content that go to make up an OU course. We will not need to repeat this activity for these types of material again.

3.3.1.2 Findings from the pre-enhancement survey

51. The outcomes of the survey⁴⁰ conducted in the first phase indicated that stakeholders do place a high value on the non-current OU learning materials, and that they perceived them to have value in all areas evaluated.
52. Headline results from the online survey revealed remarkable uniformity of opinion in terms of the value of the materials across all four dimensions; across all departments and faculties; across all respondents' primary roles within the OU community (academic, academic related, associate lecturer and secretarial and clerical); across the length of time staff had worked at the OU; and across those who were, and were not, involved in module production.
- 89.2% of respondents (501) agreed or strongly agreed with the statement that *maintaining an archive of non-current OU learning materials is important to the reputation of the OU*.
 - 75.9% of respondents thought that this should be maintained in perpetuity.
 - 90.16% of respondents (504) agreed or strongly agreed that *non-current learning materials are important to the context of the history of higher education*.

⁴⁰ <http://www.open.ac.uk/blogs/stellar/wp-content/uploads/2012/10/STELLAR-Project-Pre-enhancement-survey-report.pdf>

- 91.75% of those respondents who were involved in module production (356) agreed or strongly agreed that *when producing new OU learning material, I am likely to look to previous material, whether for inspiration or for potential reuse.*
53. When discussing the personal and professional dimension of value, it was revealed that new academics use the archive of past materials to help bed themselves in to the culture of course production at the OU; that module production was seen as an integral part of the scholarly output of academics; that there was little agreement as to how strongly asset creators felt ownership of the material produced; and that academics and course managers, particularly, had different views of the value of the archive - the former emphasising its historical value and the latter its practical value in terms of reuse. All agreed, however, that the OU produced world respected resources and that that should not be forgotten.
54. Value to the higher education and academic communities had three components: firstly, that current academics were 'standing on the shoulders of giants' and that to fail to maintain the archive was to remove the opportunity to learn from it, both pedagogically and in terms of course content; Secondly, that it has great value to future researchers and PhD students; and, thirdly, that it forms a great part of the OU's institutional legacy.
55. Investigating the archive's value to internal OU processes and cultures revealed a reliance on personal or departmental archives. It also revealed four facets of current reuse practices: firstly, that core materials in each subject are seen to be relatively stable and are often published separately from materials that change frequently, thus facilitating their reuse; secondly, a perception that many courses are removed from presentation for business or curriculum reasons and so much current course material is shelved prematurely and could be reused; Thirdly, that there is re-versioning of existing material for new and emerging markets; and, finally, that material no longer in presentation is reused in ways that may enhance the reputation of the OU.
56. When considering the financial dimension of value, interviewees focussed on four areas. Firstly on the cost-savings to be made through reuse - a complicated calculation, but one that could be made a great deal simpler if the archive were set up in such a way as to facilitate reuse; secondly, that much content could be, or is being, re-versioned in order to generate income; thirdly, that the archive could be promoted in such a way that it enhances the student experience and increases their perception of value for money; and finally, that the digitisation of past resources, while expensive, may have benefits to the academic community.

3.3.1.3 Findings from the data enhancement

57. The STELLAR project explored how the Learning Resource Metadata Initiative⁴¹ (LRMI) might be used to support the preservation of non-current learning materials. The findings from this piece of work have been shared with the wider community via the project blog⁴². It was concluded that LRMI and Schema.org represents a form of semantic technology which has the potential to expose learning resources to a wider audience via major search engines. Furthermore it will allow users to easily apply more detailed semantic queries to search results. Beyond this neither offers a semantic solution as flexible as Linked Data. For STELLAR therefore, it was decided that Linked Data would provide the semantic structure.
58. Within a previous Jisc-funded project (Lucero) the Open University had already made an attempt to transform MARC data into Linked Data using the Bibliographic Ontology (Bibo)⁴³ as the main Linked Data standard to model the catalogue records. STELLAR looked at how non-text based records could be transformed into Linked Data more specifically tailored to the asset type, using the audio-visual EBUCore profile for the university's video and audio materials. This enabled methods of metadata expression that were not available using Bibo. A much richer relationship model could be used to express the connections between these resources, so 'season', 'series' 'synopsis', 'programme', 'programme transcript', 'has subtitling' were terms that could be used to communicate and articulate these relationships much more clearly and therefore enable the user to have a much better understanding of how these materials fitted together.
59. The project also considered whether the use of MODS⁴⁴ rather than Bibo had any positive or negative impact on the ways the text-based library records could be expressed as Linked Data. Although there was no discernible advantage/disadvantage to the expression of MARC records for textbooks as MODS Linked Data records instead of Bibo, it was considerably easier to transform the records into MODS and there was a significant reduction in the time and resource required for the work. MODS did give significant advantages when it came to the expression of non-standard bibliographic materials, e.g. records for assessment materials, study guides, specimen exam papers and broadcast notes, due to the flexibility of the MODS data model.

⁴¹ <http://www.lrmi.net>

⁴² <http://www.open.ac.uk/blogs/stellar/?p=76>

⁴³ <http://bibliontology.com/>

⁴⁴ <http://www.loc.gov/standards/mods/>

60. Unfortunately, during the enhancement stage of the project, PREMIS⁴⁵ had yet to fully express its metadata profile as Linked Data, as The Library of Congress had not published the relationship elements of PREMIS in RDF by the time that the metadata profiles and relationship models were developed. The entity relationships were essential to the approach to modeling the module metadata and the lack of RDF representations of these elements in the PREMIS ontology meant that it was not possible to implement them as part of the OU study material object model (Section 8.7 below), but any relevant and available PREMIS OWL ontology terms have been integrated into our metadata profiles. The Library of Congress announced the publication of 21 new value vocabularies to be used with preservation metadata in LC's Linked Data Service for Authorities and Vocabularies⁴⁶ on June 19th 2013. 24 PREMIS value vocabularies are now available from the Library of Congress⁴⁷. Any newly published, relevant PREMIS terms will be integrated into our metadata profiles and the relationship model as part of the continuing development of the OU Digital Library.

3.3.1.4 Findings from the final evaluation

61. The full report from the evaluation phase has been published on the project blog⁴⁸. This final phase of the project has revealed that overall, academics and other stakeholders in the university do believe that the value of the selected materials was positively impacted by the application of semantic technologies. The two dimensions where the evaluation indicates the transformation of the materials has particularly increased the perceived value of the material are in the areas of value to internal processes and culture, and financial/bottom line value. The results of the questionnaire highlight that the respondents agree that semantic analysis suggested a broader range of materials than they would be presented with using a more “traditional” string search. On the basis of this, they also agreed that they would be more likely to explore non-current materials which had been enhanced in this way and such materials are more likely to be reused.
62. In the original pre-enhancement survey⁴⁹, the majority of respondents who said they had been involved in reuse of materials had found this material in a personal collection or unofficial archive. In fact only two interviewees mentioned the Library at that time as a potential source of such material. Therefore the enthusiasm for a central archive which could be effectively

⁴⁵ <http://www.loc.gov/standards/premis/>

⁴⁶ <http://id.loc.gov>

⁴⁷ <http://id.loc.gov/preservationdescriptions/>

⁴⁸ <http://www.open.ac.uk/blogs/stellar/wp-content/uploads/2013/07/STELLAR-Post-Enhancement-Survey-Report.pdf>

⁴⁹ <http://www.open.ac.uk/blogs/stellar/wp-content/uploads/2012/10/STELLAR-Project-Pre-enhancement-survey-report.pdf>

searched represents an opportunity to shift from current practice which has the disadvantages of relying on individuals and institutional memory and can suppress interdisciplinary reuse of content.

63. The majority of the participants did agree or strongly agree that the materials were easier to find than those in existing systems, indicating such a solution could potentially address current issues cited regarding the location of existing materials for reuse.
64. 83% also agreed that the OU could make savings if existing materials were enhanced as demonstrated during the evaluation. This represents an increase of around 8% when compared to the number of respondents to the pre-enhancement survey who stated they believed the OU could make savings through the increased reuse of learning materials. This increase implies that by applying linked data to the materials, making them easier to discover and giving users access to a broader range of material, STELLAR has increased the perceived financial value of the legacy content.
65. However, it was also clear that while participants felt there is financial benefit to be gained from the application of semantic technologies to digitised learning materials to increase their discoverability, there is also a cost associated with reuse of existing materials which must be factored into any calculation of savings.
- “Re the OU making savings - probably not 'savings' but better quality. Reusing materials takes time but has the potential in increase quality.”*
66. This theme was also apparent in the responses to the first benchmarking survey, where some respondents felt that potential cost savings through reuse are often overstated, particularly as reuse is not necessarily a straightforward practice.
67. The professional and personal perspectives of the value of the legacy learning materials did not seem to be strongly influenced by the application of semantic technologies. Those who do contribute to the development of modules would continue to make these available for others if the material was enhanced. However the benchmarking survey indicates this is already happening and part of the OU culture. Thus semantic technologies would not necessarily increase the desire to share material for reuse, but rather make it easier to take advantage of the content provided.
68. The evaluation indicated that participants continue to believe the legacy learning materials are valuable both to the university’s reputation and historically. They agree that making these

materials more discoverable is important, but were equally interested in exploring *which* materials are preserved in this way. Decisions regarding selection of material will inevitably influence the perception of how important and useful these materials are in the future.

69. Several face-to-face discussions in this area focused on which modules could be most valuable to preserve and enhance, and how this would contribute to the reputation of the university or the historical understanding of our own teaching and learning. Participants all agreed it was important to maintain an archive of materials, but recognised that practically it would not be possible to digitise and enhance *all* non-current modules.
69. Throughout the evaluation activities participants responded positively to the demonstration of the linked data and the improved access this provided to archived materials. They were interested in when such technologies might become commonly available in the university and could see them supporting module production processes.

“An excellent start, please keep going”

The benefits for reuse were recognised and many suggestions were made regarding what content could be preserved in future to further enrich the resources available. Such comments have been shared with colleagues in the OU and will feed into future work in this area.

3.3.1.5 Sustainability of using linked data for legacy learning materials

70. As part of the project investigations, the additional effort required to transform the selected materials into Linked Data was reviewed. The findings of this analysis are intended to indicate how sustainable this approach might be in the future and what effort and resources would need to be in place to support it.

3.3.1.5.1 Cataloguing of supplementary items

71. As described above, supplementary items were not previously catalogued, so the STELLAR project required the creation of new catalogue records for these items. In addition to this task, in some cases improvements were required to existing catalogue data to support the transformation to Linked Data.
72. The cataloguing of the supplementary items was the most time-consuming element of this work and would need to be repeated for any supplementary items digitised in the future, with

investigation and evaluation of possible methods of automating some aspects of this cataloguing required.

The effort required to catalogue supplementary items was determined by the number of items associated with each module and will always vary. However, it is estimated that as a guideline for future work it takes a Grade 5 member of staff 1 hour to catalogue 15 supplementary items.

3.3.1.5.2 Metadata

73. The metadata team undertook three elements of work for the STELLAR project; the creation of a module profile which connected the complicated web of content and metadata associated with each module, the identification and mapping of appropriate metadata schemas for the selected resource types and the extraction of data relating to the non-current modules from university systems. The majority of this work can be considered a one-off task. If modules comprising the same formats of materials are digitised and preserved in the future, the metadata mapping has already been completed and will not need to be repeated. The extensive nature of the metadata means that in future it should be relatively easy to add new content into OUDL for the resource types STELLAR has focussed on. Therefore no additional effort would be required. Similarly, the data extraction will not need to be repeated on a similar scale in the future - a scheduled report could be run to update the data held with details about newly retired modules. If a new format of material is introduced in the future then a new metadata schema would be needed.

Development of module profile – this task will be required in the future when a new collection featuring new resource types is added. It is estimated that adding one new collection featuring new resource types would involve an effort of 39 hours from a Grade 7 Metadata Specialist. This is over and above the 19 hours of effort per month required for other routine tasks relating to adding new collections.

3.3.1.5.3 OUDL development

74. The module profile, module metadata and relationship model developed by the metadata team were ingested into the OUDL. In addition, work in this area focussed on surfacing new content in the search and/or user interface. This required considerable effort and was an iterative activity.
75. The scanned data had to be matched to the catalogue records, which proved to be time-consuming due to the unpredictable formats of the data. In theory once this has been done, the activity can be repeated using the same scripts. Adding the semantic linking data also required initial expenditure of effort to learn what was required. However, actually

implementing it was relatively low effort because the data was already organised. The most intensive linking was that of linking the book items to the page items (this was not semantic linking) so that the actual adding of RDF was relatively quick.

76. RDF has now been built into OUDL and the relationships have been modelled within the system. In the future the process of creating linked data for similar content should take little, or negligible amounts of time, although there will be further enhancements. However, time saved in this activity will be spent on building more applications and interfaces based on the linked data work.

3.3.2 Lessons learnt

3.3.2.1 Engagement with the community

77. There was an assumption the high levels of engagement achieved with the pre-enhancement survey could be maintained throughout the lifetime of the project. A significant number of individuals (approximately 170) involved in the first survey expressed an interest in remaining involved in the project going forward. The intention was to focus ongoing communication and involvement activities on this group.
78. However while many colleagues were keen to express their views on the value of legacy learning materials, they were less enthusiastic to become involved with the subsequent evaluation of the transformation of the selected collections. As a result the evaluation was conducted with a much smaller group of stakeholders than the pre-enhancement survey. Once people did attend a workshop or demonstration they tended to be highly engaged, actively participated in discussion and offered feedback. They often volunteered to share what they had seen with their colleagues. The challenge was encouraging them to become involved in the first place. It is concluded that there are two possible reasons for this:
79. There were several months between the pre-enhancement survey and evaluation activities. Although regular communications had been sent to the individuals who had expressed an ongoing interest in the project they did not respond in significant numbers to the invitations to take part in the evaluation. Engagement appeared to have declined over time.
80. While the pre-enhancement survey involved completing a short online survey, more effort was involved in participating in the evaluation. This required attendance at a face-to-face event or completing a series of online activities from one's desk. It may be that many people felt they could not commit the time this would require. If engagement was not high when the invitation

to participate was received, it is likely that they were unwilling or unable to consider dedicating the time in their schedule it would require to take part.

81. In retrospect the evaluation phase could have been launched in a similar fashion to the first survey; as if starting from a baseline of low awareness of and engagement with the project. If stakeholders were supportive of the project and aware of its goals and achievements, then they would be more likely to commit the time to participating in the review of its outputs.

3.3.2.2 Data enhancement

82. The decision was taken to include the supplementary items in the digitisation of a whole module. These types of learning materials had not been digitised before and required additional cataloguing time. It became apparent that to prepare a complete module for ingestion into OUDL, including digitisation and cataloguing of these supplementary items, requires considerable effort.
83. The classification of some types of OU study material as 'supplementary' is due to historic, print-centric publishing processes within the university. This labelling of all materials that were not considered as 'main texts' has had a considerable impact upon how these materials have been managed and treated. Supplementary materials include study guides for the module and the individual units, readings and off-prints, assessment materials, workbooks and broadcast notes⁵⁰. The inclusion of supplementary items has highlighted the fact that some of these resources can have high levels of contextual value to the assets they relate to. For example, Broadcast Notes would be extremely helpful to our understanding of module audio and video assets, and if they were more explicitly linked, they could make it easier for users to find relevant audio-visual resources.
84. The existence of an XCRI validator, together with the corpus of implementations⁵¹ greatly facilitated modelling the course records in XCRI. In particular previous work on the OU's own XCRI feed⁵² was utilised together with documentation from the XCRI Knowledge Base⁵³.
85. There were not any suitable academic metadata standards for describing our resources. Although we were able to use XCRI for describing the modules offered by the university there

⁵⁰ <http://www.open.ac.uk/blogs/stellar/wp-content/uploads/2013/07/OU DL-STELLAR-metadata-implementation.docx>

⁵¹ http://www.xcri.org/wiki/index.php/XCRI_Feeds

⁵² <http://www3.open.ac.uk/study/feeds/ou-xcri-cap.xml>

⁵³ <http://www.xcri.co.uk/>.

were no suitable standards for describing our assets. IEEE LOM⁵⁴ has no RDF representation. LRMI is able to represent learning resources at a very broad level and will allow us to share those resources with other organisations but it is not currently suitable for representing the complexity of our assets to meet our internal requirements and needs.

86. There is no suitable vocabulary for describing learning resource types currently available. LRMI is not specifying or advocating the use of specific vocabularies at this stage in its development. We reviewed vocabularies used by other organisations (most have been developed to meet local needs), and created our own list mapped to existing vocabularies⁵⁵.
87. A major part of the learning was gaining a very precise knowledge of exactly the extent and limitations of Fedora's use of RDF and the implications of these. Fedora's inability to nest RDF statements was a major shortcoming and enforced a re-write of the object relationship modelling.
88. Fedora's enforcement of relationships only between Fedora objects and/or Fedora datastreams was less significant in reducing functionality than anticipated. It proved both possible and desirable to reference objects and models which were 'legal' Fedora syntax terms without enforcing actual existence of the referenced objects and we used a series of 'models' to give precision to relationships which Fedora's RELS-EXT⁵⁶ datastream was not capable of expressing: for example,

```
programme 'A' has transcript 'B'

was expressed as

program A has 'B'          and

'B' hasModel 'transcript'.
```

89. It was not possible to add RDF content external to Fedora to its internal triplestore. The significant shortcomings of this for everyday operations were not initially appreciated, enforcing wrangling content in to legal FOXML that also had to be un-nested RDF.
90. Developing XML-based metadata is not difficult and developing RDF versions of metadata profiles is getting easier to achieve as more standards come on-board, but developing a repository for both the XML and RDF versions of the metadata schema is significantly more

⁵⁴ http://ltsc.ieee.org/wg12/files/LOM_1484_12_1_v1_Final_Draft.pdf

⁵⁵ <http://www.open.ac.uk/blogs/stellar/wp-content/uploads/2013/07/OU DL-learning-resource-types.xlsx>

⁵⁶ <http://fedora-commons.org/download/2.2/userdocs/digitalobjects/introRelsExt.html>

demanding and time-consuming. Many metadata schemas are using slightly different models for their RDF other than the XML schema, this makes it difficult to house both representations. In some standards the difference between the XML and RDF is quite significant. This was the case with EBUCore RDF ontology, which uses the concept of business objects, editorial objects and media objects that are not part of the XML representation and therefore made it quite difficult to map to some parts of our existing XML-based schema.

91. One foreseen outcome (because of previous work with Lucero) was that the presentation of content in linked data format would show in an unmerciful light any shortcomings in data entry and highlight artefacts in previous cataloguing.

3.4 Immediate Impact

92. The findings from the initial survey conducted as part of the STELLAR project have contributed to the articulation of the value placed on OU non-current learning materials within the institution, and provided a strong evidence base for this. The survey findings have been widely disseminated within the OU and used as reference for other projects. For example a recent study looking at user requirements of academics and other staff when searching for materials to re-use in OU courses (MacGill 2013 'User requirements study of academics and other staff when searching for material (found content) to re-use in Open University (OU) courses', *unpublished internal OU report*)
93. The development of the prototype tool has allowed the project team to effectively demonstrate the benefits of using semantic technologies to the wider OU community. This will continue to support an ongoing discussion within the OU about improvements to existing search facilities and will inform future projects and decision-making.
94. Recommendations from the pre-enhancement survey have also been fed into the OU Digital Preservation Policy and criteria for selection of legacy learning material to preserve (see sections 8.8 and 8.9 for the draft policy and selection criteria). The policy outlines the scope of and the key principles underpinning the university's approach to digital preservation.
95. For the first time the Open University has acquired a single accessible, simple to understand and comprehensive list of all the courses and all the presentations of these courses since the university's inception without requiring extensive technical intervention. The OU has access to a full record of all the resources related to the sample courses (text, audio, video and web) in a single place, available through a single interface.

96. There is also now a location in which to sustain a permanent record of new course data (currently updated and released daily as xcri-cap feeds).
97. The OU also has a more useful way of relating resources together that is especially valuable when dealing with complex multi-part objects such as modules and programmes. For example, It is not easily possible to make explicit in a relational database that a programme has three different master broadcast recordings, 11 associated manuscripts, is one of 32 programmes related to a given rights contract bundle, has subtitles etc. However, to model this precisely – and extend it further – is easily expressed in linked data.
98. A significant finding of the project was the substantial additional value to be gained from the existing corpus of OU linked data with minimal developer effort required to display contextual information from open data iTunesU and Open Learn resources within the context of the digital library resources. This proved of significant value, even without semantic enhancement.
99. The Open University library catalogue contains records go back to the university's first course presentations but they have not been catalogued consistently and were not clearly badged as Open University created content. The lessons learnt from migrating content into the OU Digital Library from STELLAR are now being used as a template for the migration of all OU study material records from the current library catalogue to the OUDL.

3.5 Future Impact

100. The availability of this information within OUDL will be a significant resource for all those involved in course production in the OU. This will impact on future cataloguing of course resources. The accessibility of the content rather than just the metadata should contribute to making content easier to find. However, the detailed cataloguing of supplementary materials undertaken for the project is not sustainable and will need to be revised.
101. It became apparent from presenting course contents in a single interface for the first time that minimal cataloguing of course content, by unit number and type of resource is invaluable in making the content easy to discover and re-use. Presenting lists of course materials by physical type (e.g. 'printed material') gave unmanageably long lists (often over 200 items) of content that require further categorisation to be useful.
102. The availability of this content outside OUDL is a significant enhancement to OU's total linked data volume and a range of further applications built across different linked data and search

index information sources is being planned. There seems to be a developing 'critical mass' of linked data at the OU sufficient to make it worthwhile to continue to join up other data sources.

103. STELLAR development opens an opportunity to use a more typical 'Fedora way of working' with content models⁵⁷ to provide service definitions and dissemination of objects. This is a valuable alternative to the current methodology, working with REST API calls, but will require additional investment to develop.
104. It is anticipated that SPARQL queries will be developed to pull in content: both directly related to the digital library but not part of it (such as still images) and from other OU data sources (such as Open Learn).
105. It is less obviously sustainable to utilise the more semantic elements of linked data, in a production setting, as the sense engine utilised by KMI in the STELLAR proof-of-concept tool is currently deployed on an experimental rather than a production basis. Two further developments would be required to make it useful in a production setting. Changes would be required to allow the tool to:
 - be capable of correction or addition of new terms as well as adjustment of existing terms
 - intelligently collapse multiple results for content from a single object (e.g. different pages from a book)

4 Conclusions

4.1 General

106. The methodology used within the Programme Strand with a pre-enhancement survey, followed by some enhancement and a follow up post-enhancement stage has proved to be an extremely useful approach. The pre-enhancement stage has, for the first time, allowed us to find ways of articulating the value of the collection in a coherent way. This report has already been useful beyond the project as a way of briefing key University stakeholders in the value of that material. It continues to inform policy in this area, not just for the Digital Library but for wider conversations about archiving and re-use of material.

4.2 Value of legacy learning materials

⁵⁷ <https://wiki.duraspace.org/display/FEDORA36/CMA+Construction+Guide>

107. When considering the objectives of the STELLAR project, it can be concluded that the stakeholders in The Open University do place value on legacy learning materials. This value takes a number of forms, but underpins the belief that such materials should be preserved. Preservation will enhance the reputation of the university, contribute to the understanding of the history of both the university and HE in general, provide academics and students with the opportunity to learn from this rich resource and support future module production by facilitating reuse of content.
108. The findings of the STELLAR research and that undertaken in the MacGill 2013 report '*User requirements study of academics and other staff when searching for material (found content) to re-use in Open University (OU) courses*' indicate that OU stakeholders involved in module production value the opportunity to refer to or reuse non-current materials:

"We have, at the OU, an amazing resource – if only it could be better organised and easier to access. I think, with better access, and wider knowledge of the existence of these resources, we could go a long way to saving production costs". (Respondent to STELLAR capturing perceptions survey)

109. Stakeholders however continue to highlight the difficulties they encounter in accessing this valuable material, and therefore there is clearly an appetite for improved searching of assets and the ability to link to other datasets within the OU.
110. It should be noted that for the purposes of the STELLAR project the user population for these learning materials comprised internal stakeholders, as this material is mainly available to OU staff. OU students and the wider HE community may have different perspectives of the value of that material.

4.3 Application of semantic technologies to legacy learning materials

111. The evaluation conducted following the transformation of the selected materials indicated that those involved believe the value of the materials has been added to by taking this approach⁵⁸. When demonstrating the application of linked data to the selected learning materials to groups of users within the OU community, they confirmed they believed that this type of technology enabled them to access a wider range of materials. The materials were more discoverable and

⁵⁸ <http://www.open.ac.uk/blogs/stellar/wp-content/uploads/2013/07/STELLAR-Post-Enhancement-Survey-Report.pdf>

as a result they were more likely to refer to materials enhanced in this way during module production. This they also believed could lead to cost savings for the university.

112. A majority of those questioned agreed or strongly agreed that:
- *Enhancing materials in this way makes them easier to reuse or refer to for inspiration during module production*
 - *Existing materials enhanced in this way are more likely to be referred to during module production than those preserved in existing OU systems*

4.4 Sustainability of using semantic technologies for such collections

113. Matching digitised content to cataloguing records for the content proved problematic, despite the existence of established naming conventions. Content data was essentially no more predictable for the last batch of content than it had been for the first, requiring continuous re-writing of 'matching' scripts. Resolution of this is required to make the project's approach sustainable.
114. We found that the current educational ontologies available were not sufficient to represent the complexity of the data we hold in the Open University and that further development of both XCRI and LRMI would significantly improve this situation.
115. Additional cataloguing of content was expensive in time and money and in future a variety of methods to enable the inclusion of previously uncatalogued content into the repository should be investigated. A number of ways of automating basic cataloguing tasks should be evaluated.
116. Enhancing items with RDF relationships was a technically trivial undertaking once a profile capable of transformation had been established and accounted for less than one fifth of the developer's time on this project. Fedora's flexibility as a content store for a wide variety of content and its inbuilt RDF functionality greatly facilitated this, once its limitations were well understood. Conversely, from a metadata perspective, creating RDF relationships for the Fedora repository was difficult and frustrating. The RDF nesting limitations of Fedora heavily impacted on the amount of resource the RDF relationship model took to develop and resulted in the model being much less rich and detailed than it had the potential to be.

117. Further linked data development within OU digital library would be facilitated by re-configuring the Fedora repository to utilise an external triplestore, rather than (the restricted functionality of) an internal one.
118. Metadata XML schemas and their RDF representations often diverge and in some cases this divergence was quite significant. This is often an unavoidable consequence of the process of developing RDF schemas/Linked Data ontologies, but it does pose problems for many repositories that wish to utilise both XML and Linked Data, as it is frequently quite awkward and time-consuming to keep the mappings for both methods of representation aligned. Effort concentrated on developing innovative and simpler ways of enabling the functionality these schemas give us would be invaluable.

5 Recommendations

5.1 Recommendations for future development

1	Further work to establish relationships between discontinued courses and successor current courses might enable related resources to be utilised more fully. Currently only direct descendent relationships are recorded ('course X has successor course Y') but deploying a fuller ontology of branching and qualification paths might facilitate discovery
2	Further development of the pathways to qualifications within XCRI is currently problematic as OU modules provide enormously diverse pathways to qualifications that cannot be satisfactorily expressed in XCRI. Additional metadata fields (and constraints on combinations) would be required to develop this in XCRI.
3	The OU may wish to develop a process for harvesting new XCRI data from the existing continuously updated live XCRI feed for archiving
4	Further enrichment of course data within the Digital Library should be undertaken with data available from other, non-discoverable sources, selected by the University Archivist.
5	Currently matching the scanned content to catalogue records in the OUDL is time-consuming and not easily repeatable. More consistent use of naming conventions for scanning and cataloguing is required to reduce effort in this area.
6	The OU Digital Library development should move towards a more linked data approach to develop library resources and rely less upon Fedora's REST API for resource discovery.

7.	The OU should continue to explore the potential of semantic web technologies as a tool to search across all learning and teaching assets, not just those preserved within the Digital Library.
8	<p>The OU Digital Library should look at a number of methods for increasing efficiency of cataloguing supplementary items:</p> <ul style="list-style-type: none"> ○ Providing resolvable identification and classification, as outlined, to establish minimal “what-type-of-item” and “which item” information. To identify where it would be sufficient to be working with full-text indexing to make content self-describing; and using this, together with RDF to elucidate content items relationships with each other, make content discoverable via search and present it in context. ○ Identify which types of supplementary items give the most contextual value to other module resources and determine whether these resources need to be treated differently to other supplementary item types and where necessary develop new metadata profiles to capture this information and allow us to identify the relationships between these materials. ○ Identifying what tasks need cataloguing expertise to achieve.
9	The Open University should phase out the use of the term ‘supplementary item’ which devalues and generalises this body of diverse learning resources. Instead, these materials should be given distinct identities and a clearer understanding of their value within the OU study material corpus should be developed.
10.	The OU should consider whether follow up work should be done to compare the perceptions of the value of the non-current learning materials that has been identified by internal stakeholders with perceptions of other groups, particularly OU students, but also the wider HE community. This would help to inform future policymaking and decisions about which material should be prioritised to be made openly available.

5.2 Recommendations relevant to the wider community

1	An update to XCRI validator would make it even more valuable tool than it currently is.
2.	Extending the range of triplestores to which Fedora can attach and change the default

	triplestore to a more obvious default as Mulgara is not a popular default triplestore today.
3.	Developing Fedora triplestores to more readily have 'external' input (either by defaulting to an external triplestore at install time or by refining the means by which the current internal Fedora triplestore can be updated). Alternatively, the flexibility of adding non-fedora content to an internal triplestore would also be useful.
4.	Extend the range of content types (ie beyond Dublin Core and RELS) which are transformed into RDF/triples on ingest within Fedora: for example to enable automated transformation of XCRI-CAP into RDF.
5.	The pre-enhancement, enhancement, post enhancement methodology is a very useful process and a technique that could be widely used across the sector.

5.3 Recommendations for Jisc

1.	<p>UK Higher and distance education needs to continue to have a clear voice in the LRMI community and ensure that the schema continues to evolve to meet the needs of learning resource creators and consumers outside of the United States. Clear recommendations regarding how UK universities should use LRMI to describe learning resources should be developed, which encompass:</p> <ul style="list-style-type: none"> the use of 'educationalAlignment' and its sub-property 'educationalFramework', particularly in relation to 'educationalSubject' and 'educationalLevel' the development of a set of controlled terms for 'interactivityType' that are meaningful to a wide audience of users a set of controlled terms for the 'educationalUse' element <p>These terms do not need to be rigidly fixed and could be co-ordinated through the use of a wiki or other social medium to record and co-ordinate the values but <i>some</i> form of co-ordination and management of appropriate terms is recommended</p>
2.	Develop the XCRI schema to better meet the needs of online and distance education, this could be achieved through making it easier to describe learning resource providers that are not based in a physical campus environment.
3.	Students at The Open University study a number of modules to gain a qualification with us. A

	module can run for more than a year, making it is necessary for us to assign educational levels to each module we teach, but this model doesn't fit with the way XCRI is currently designed. We would recommend that Jisc looks at enable the assignment of educational level at the course level rather than just to a qualification. This would enable us to more comprehensively adopt XCRI for all Open University offerings.
4.	Jisc funding for projects of this nature are immensely valuable at both an institutional and community level. Jisc should continue to be a source of funding for this type of exploration as it is critical to build knowledge in the sector.
5.	The programme methodology of using a pre-enhancement stage, enhancement, followed by a post-enhancement stage has worked well for this project. Jisc should use this methodology, where appropriate, for future activities.

6 Implications for the future

6.1 *Internal implications*

119. Although there are benefits to digitising an entire module as this preserves the whole experience, there are costs associated with this, particularly for older modules. Additional digitisation and cataloguing have been required, as items such as the print supplementary items are currently not preserved digitally and do not have catalogue records. Therefore extra resource would be required to complete these additional tasks if this approach is taken in future. In addition the level of cataloguing required for smaller items such as for the supplementary items can be significant, but some of these items may not be as reusable as the study materials and AV materials. This project has been a first attempt to catalogue these types of items and more work needs to be done to assess which items can be catalogued and indexed with little human intervention and which need further investigation to identify the data and relationships that needs to be captured, to enable the extraction of maximum value from the resources. This may mean further refinement of the metadata profiles we have developed for these types of assets.
120. The learning about Fedora's RDF capabilities has been extensive and this will be useful in the future. However, as it is a different paradigm of development, it will be deployed in new and developing areas of content such as the creation of multiple overlapping collections of material. It is already being put to use to show relationships between data that databases struggle to make clear. A custom ontology would enable us to be more precise about the relations between objects as expressed in Fedora, but would take significant resource to develop.
121. The STELLAR search tool has started to identify the value of semantic search, where a user is searching for 'meaning' rather than simple string matching. Future exploration in this area may prove valuable as it is likely to be of considerable value to searchers looking for found content for re-use. However the effort needed to catalogue and transform existing content into RDF is not insignificant and is a barrier.

6.2 *Implications for wider community*

122. Although a pioneering RDF-capable repository, Fedora's linked data capabilities have not kept pace with current developments and additional linked data functionality within Fedora would be extremely valuable for the wider Fedora community.

123. Utilising existing linked data resources is an easy way to augment newly created linked data and this project has found that the benefits of doing this seem to be cumulative: it is proportionately more effective and useful to extend linked data resources the more associated data sources that already exist.

6.3 *Ongoing availability of project outputs*

124. Internally the STELLAR proof-of concept tool will remain available as a means of demonstrating how Linked Data can be applied to enhance the discoverability of content and specifically learning materials. This will continue to be used to stimulate and support debate within the institution regarding development of the digital library and improvements to searching for content across university repositories.
125. The STELLAR project blog will be maintained for five years following the project closure, making available an overview of the project approach and providing access to the key reports and outputs from the project. The project blog will be archived in the OU Digital Library.
126. Key reports will also be deposited in JORUM and in the OU Digital Library.

7 References

Currall, J., McKinney, P. 'eSpida handbook – Expressing project costs and benefits in a systematic way for investment in information and IT' 2007 *University of Glasgow*
http://dspace.gla.ac.uk:8080/bitstream/1905/691/1/espida_handbook_web.pdf (viewed 28 June 2013)

d'Aquin, M., Allocca, C. and Collins, T. 2012 'DiscOU: A flexible discovery engine for open educational resources using semantic indexing and relationship summaries', *Demo at International Semantic Web Conference, ISWC 2012* <http://data.open.ac.uk/applications/iswc2012-demo.pdf> (viewed 28 June 2013)

d'Aquin, M. (2012) 'Linked data for open and distance learning', *Common Wealth of Learning*
<http://www.col.org/resources/publications/Pages/detail.aspx?PID=420> (viewed 28 June 2013)

Donaldson, D. & Conway, P 2010, 'Implementing PREMIS: a case study of the Florida Digital Archive', *Library Hi Tech*, 28, 2, pp. 273-289, Social Sciences Citation Index, EBSCOhost, (viewed 28 June 2013)

Kaplan R.S., and D.P. Norton. 1992 'The balanced scorecard - measures that drive performance' *Harvard Business Review* 70: 58-63 (viewed 28 June 2013)

Lagoze, C, Payette, S, Shin, E, & Wilper, C 2006, 'Fedora: an architecture for complex objects and their relationships', *International Journal On Digital Libraries*, 6, 2, pp. 124-138, Academic Search Complete, EBSCOhost, (viewed 28 June 2013). DOI: 10.1007/s00799-005-0130-3

MacGill L., Falconer, I. (2013) "User requirements study of academics and other staff when searching for material (found content) to re-use in Open University (OU) courses' *Lou MacGill Consulting* (Unpublished consultancy report)

McKinney, P., 2005. 'espida and Sustainable Digital Preservation'. *Archives, Records Management & Conservation Society of Archivists* 193 <http://hdl.handle.net/1905/443> (viewed 28 June 2013)

Pennock, M., 2007, 'Supporting institutional digital preservation & asset management: A summary of the JISC DPAM programme synthesis', *New Review Of Information Networking*, 13, 2, pp. 119-132, Academic Search Complete, EBSCOhost, (viewed 28 June 2013) DOI: 10.1080/13614570802085962

'Sustainable economics for a digital planet: Ensuring long-term access to digital information – Final report of the Blue Ribbon Task Force on sustainable digital preservation and access' 2010 Blue Ribbon Task Force http://brtf.sdsc.edu/biblio/BRTF_Final_Report.pdf (viewed 28 June 2013)

Tanner S., 2008 'Open University Footsteps project – Investigating the feasibility of digitising and processing the Open University course materials archive' *KCL* (Unpublished consultancy report)

Zablith, F., d'Aquin, M., Brown, S. and Green-Hughes, L. (2011) 'Consuming linked data within a large educational organization', *Workshop: Second International Workshop on Consuming Linked Data (COLD) at International Semantic Web Conference (ISWC), Bonn, Germany*
<http://fouad.zablith.org/docs/ISWCConsumingLinkedData2011.pdf> (viewed 28 June 2013)

Metadata references

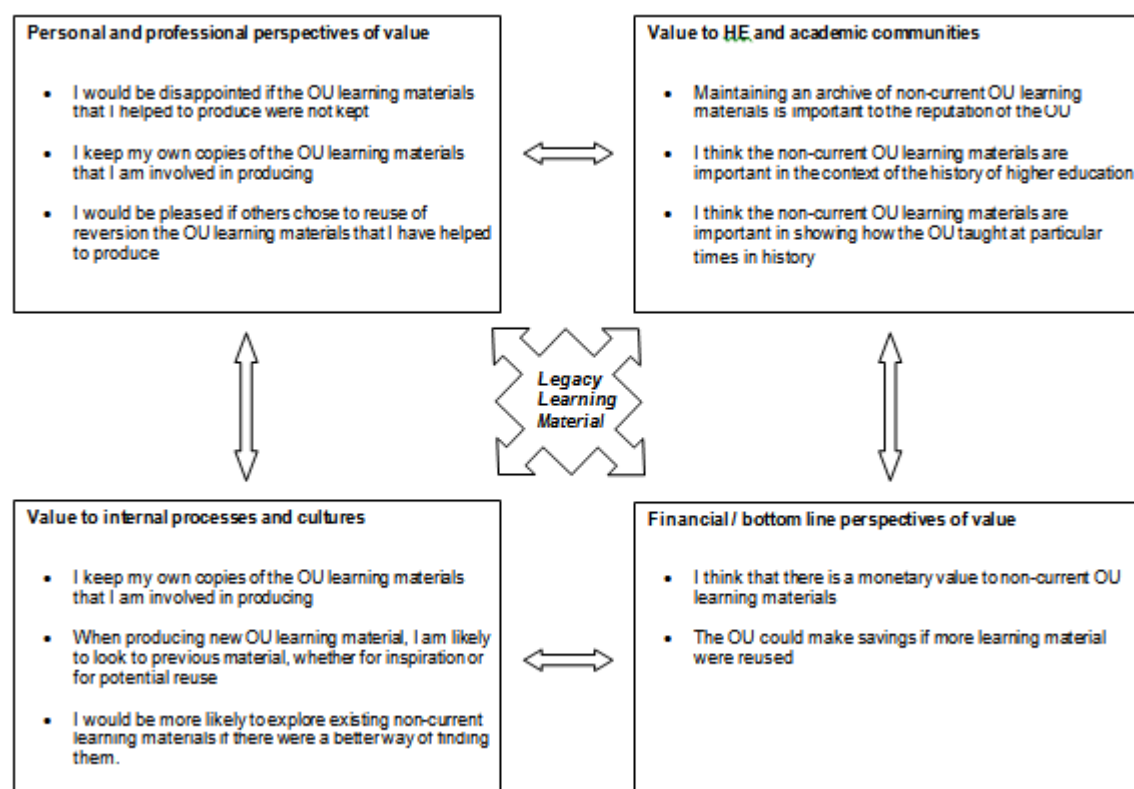
Bibo <http://bibliontology.com/>
EBUCore <http://www.ebu.ch/metadata/ontologies/ebucore/>
Learning Resource Metadata Initiative <http://www.lrmi.net/>
LOM http://ltsc.ieee.org/wg12/files/LOM_1484_12_1_v1_Final_Draft.pdf
MODS <http://www.loc.gov/standards/mods/modsrdf/primer.html>

Project Identifier:
Version: 1.0
Contact: Sarah Brown
Date: 17/09/13

PBCore <http://pbcore.org/index.php>
PREMIS <http://www.loc.gov/standards/premis/>
PREMIS vocabularies <http://id.loc.gov/preservationdescriptions/>
XCRI <http://www.xcri.co.uk/>

8 Appendices

8.1 *Balanced scorecard*



8.2 Pre-enhancement survey questionnaire

EXPLORING THE VALUE OF OUR LEARNING MATERIALS

Library Services is undertaking this survey as part of a JISC-funded project⁽¹⁾ to find out how academics and key stakeholders in the university value the archive of OU learning materials. The results will provide us with critical information about what you value and what should be preserved for the future.

What should happen to OU learning materials when they are withdrawn from presentation? Is it important to you that they are kept? If so, for how long? How can we best ensure that they are accessible over time?

We look forward to hearing your views. If you have any questions regarding this survey or the project, please email: stellar@open.ac.uk or visit the blog at <http://www.open.ac.uk/blogs/stellar>

⁽¹⁾ *Stellar – Semantic Technologies Enhancing the Lifecycle of Learning Resources*
<http://www.open.ac.uk/blogs/stellar>

Terminology:

For the purposes of this survey OU learning materials includes any OU produced materials provided to students as part of a module. Non-current learning material refers to study material which belongs to any module which is no longer in presentation.

Data Protection Information:

The data you provide will be used for research and quality improvement purposes and the raw data will be seen and processed only by The Open University staff and its agents. This project is administered under the OU's general data protection policy guidelines.

Please return this questionnaire, via internal mail, to:

**Liz Mallett
Library Services**

q1 - Which Faculty/Unit do you work in?

q2 - Is your primary role at the OU...

- ☐ Academic?
☐ Academic related?
☐ Associate Lecturer?
☐ Secretarial/Clerical?

q3 - How long have you worked at the OU?

- ☐ Fewer than 3 years
☐ 3-10 years
☐ Over 10 years

q4 - Are you, or have you ever been, involved in module production at the OU?

- ☐ Yes
☐ No

q5 - To what extent do you agree or disagree with the following statements?

(Please make one selection for each statement)

	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
Maintaining an archive of non-current OU learning materials is important to the reputation of the OU	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I think that there is a monetary value to non-current OU learning materials	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I think the non-current OU learning materials are important in the context of the history of Higher Education	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I think the non-current OU learning materials are important in showing how the OU taught at particular times in history	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

If you are, or have ever been, involved in module production, please answer all remaining questions.

If you have not, and have never been, please skip to question 7

q6 - To what extent do you agree or disagree with the following statements?

(Please make one selection for each statement)

	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
I would be disappointed if the OU learning materials that I helped to produce were not kept	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I keep my own copies of the OU learning materials that I am involved in producing	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I would be pleased if others chose to re-use or re-version the OU learning materials that I have helped to produce	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
When producing new OU learning material, I am likely to look to previous material, whether for inspiration or for potential re-use	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

q7 - To what extent do you agree or disagree with the following statements?

(Please make one selection for each statement)

	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
I would be more likely to explore existing non-current OU learning materials if there were a better way of finding them	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The OU could make savings if more learning material were re-used	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

q8 - In summary, I think that the OU should keep non-current learning materials.

- ☐ In perpetuity
- ☐ For 5-10 years
- ☐ For up to 5 years
- ☐ Not at all

q9 - If you would like to give further thoughts/comments about archiving non-current OU learning materials and/or their re-use, please use this space:

q10 - As part of the next stage of this study, we would like to carry out some short interviews with OU staff members

Would you be willing to be contacted about this?

- ☐ Yes
☐ No

If you answered yes to question 10 or 11, please provide your contact details:

q11 - We are also setting up a panel to discuss aspects of this project as it continues.

Would you be interested in being part of this?

- ☐ Yes
☐ No

Name:

Email:

Ext:

q13 - Would you like to receive a summary of the responses to this survey?

- ☐ Yes
☐ No

q14 – If so, what address would you like us to email the summary to?

Complete – Many thanks for taking the time to complete this survey

Your feedback is invaluable to us and will help us in future policy formation. You will now be directed to the Open University website.

8.3 Post-enhancement survey

Thank you for taking the time to review the enhancements the STELLAR project has made to the selected OU learning materials. Your feedback on the value of these enhancements will form part of our evaluation and feed into our final report.

Please complete this short questionnaire to provide us with your feedback and comments. This should take no longer than 10 minutes.

1. Which Faculty / Unit do you work in?
2. Is your primary role at the OU:
 - Academic?
 - Academic related?
 - Associate Lecturer?
 - Secretarial/Clerical?
3. How long have you worked at the OU?
 - Fewer than 3 years
 - 3 – 10 years
 - Over 10 years
4. Are you, or have you ever been, involved in module production at the OU?
 - Yes
 - No
5. How have you participated in the STELLAR evaluation?
 - Attended a face-to-face
 - workshop
 - Attended a Unit/Faculty demonstration
 - Completed an online review

Based on your experience today, to what extent do you agree or disagree with the following statements?

6. Based on your experience today, to what extent do you agree or disagree with the following statements?

	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
The semantic analysis suggested existing materials would not have found using a traditional search.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I would be more likely to explore existing materials if I knew they had been enhanced in this way	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Enhancing materials in this way makes them easier to reuse or refer to for inspiration during module production	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I think others would be more likely to	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

refer to materials I have helped to produce if they were enhanced in this way.					
I would make materials I have helped to produce available for similar enhancement	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Making noncurrent module materials more easily discoverable through semantic technologies is important to the reputation of the OU.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Making noncurrent module materials more easily discoverable through semantic technologies adds to the understanding of the history of Higher Education	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Existing materials enhanced in this way are more likely to be referred to during module production than those preserved in existing OU systems	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The OU could make savings if existing materials were enhanced to make them more discoverable	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

7. If you would like to give further thoughts or comments regarding the enhancement of Non-current OU learning materials, please use this space:

Thank you for taking part in the STELLAR evaluation and completing this questionnaire.

8.4 Fedora XCRI module record for S100

Edited version to remove repeated elements

```
= <foxml:digitalObject VERSION="1.1" PID="xcri:S100"
  xmlns:foxml="info:fedora/fedora-system:def/foxml#"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="info:fedora/fedora-system:def/foxml#
    http://www.fedora.info/definitions/1/0/foxml1-1.xsd">
-
= <foxml:objectProperties>
  <foxml:property NAME="info:fedora/fedora-system:def/model#state"
    VALUE="Active" />
  <foxml:property NAME="info:fedora/fedora-system:def/model#label"
    VALUE="xcri:S100" />
  <foxml:property NAME="info:fedora/fedora-system:def/model#ownerId"
    VALUE="fedoraAdmin" />
  <foxml:property NAME="info:fedora/fedora-system:def/model#createdDate"
    VALUE="2013-04-22T10:16:29.606Z" />
  <foxml:property NAME="info:fedora/fedora-system:def/view#lastModifiedDate"
    VALUE="2013-04-22T10:16:29.606Z" />
</foxml:objectProperties>
-
= <foxml:datastream ID="AUDIT" STATE="A" CONTROL_GROUP="X"
  VERSIONABLE="false">
-
= <foxml:datastreamVersion ID="AUDIT.0" LABEL="Audit Trail for this object"
  CREATED="2013-04-22T10:16:29.606Z" MIMETYPE="text/xml"
  FORMAT_URI="info:fedora/fedora-system:format/xml.fedora.audit">
-
= <foxml:xmlContent>
-
+ <audit:auditTrail xmlns:audit="info:fedora/fedora-system:def/audit#">
  </foxml:xmlContent>
  </foxml:datastreamVersion>
  </foxml:datastream>
-
= <foxml:datastream ID="DC" STATE="A" CONTROL_GROUP="X" VERSIONABLE="true">
-
= <foxml:datastreamVersion ID="DC1.0" LABEL="Dublin Core Record for this object"
  CREATED="2013-04-22T10:16:29.629Z" MIMETYPE="text/xml"
  FORMAT_URI="http://www.openarchives.org/OAI/2.0/oai_dc/" SIZE="567">
-
= <foxml:xmlContent>
-
= <oai_dc:dc xmlns:oai_dc="http://www.openarchives.org/OAI/2.0/oai_dc/"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai_dc/
    http://www.openarchives.org/OAI/2.0/oai_dc.xsd">
```

```
<dc:title>Science: a foundation course</dc:title>
<dc:description />
<dc:contributor>Science Faculty</dc:contributor>
<dc:format />
<dc:identifier>xcri:S100</dc:identifier>
<dc:identifier>S100</dc:identifier>
<dc:source />
  </oai_dc:dc>
  </foxml:xmlContent>
  </foxml:datastreamVersion>
  </foxml:datastream>
-
- <foxml:datastream ID="XCRI" STATE="A" CONTROL_GROUP="X" VERSIONABLE="true">
-
- <foxml:datastreamVersion ID="XCRI.0" LABEL="XCRI" CREATED="2013-04-
  22T10:16:29.629Z" MIMETYPE="text/xml" SIZE="11303">
-
- <foxml:xmlContent>
-
- <xcri:catalog xmlns:xcri="http://xcri.org/profiles/1.2/catalog"
  xmlns:xcriTerms="http://xcri.org/profiles/catalog/terms"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns:xhtml="http://www.w3.org/1999/xhtml"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:dcterms="http://purl.org/dc/terms/"
  xmlns:credit="http://purl.org/net/cm" xmlns:mlo="http://purl.org/net/mlo"
  xmlns:courseDataProgramme="http://xcri.co.uk"
  xmlns:soton="http://id.southampton.ac.uk/datasets/jacs"
  xsi:schemaLocation="http://xcri.org/profiles/1.2/catalog
    http://www.xcri.co.uk/bindings/xcri_cap_1_2.xsd
    http://xcri.org/profiles/1.2/catalog/terms
    http://www.xcri.co.uk/bindings/xcri_cap_terms_1_2.xsd http://xcri.co.uk
    http://www.xcri.co.uk/bindings/coursedataprogramme.xsd" generated="2013-
    02-22T11:19:02Z">
-
- <xcri:provider>
-
- <xcri:course>
  <dc:contributor>Science Faculty</dc:contributor>
  <dc:date xsi:type="dcterms:W3CDTF">1971-1978</dc:date>
  <dc:date>Years in presentation: 8</dc:date>
  <dc:description xcri:type="providedResource" />
  <dc:description xsi:type="xcriTerms:studyHours">Approximately 600 hours of
    study</dc:description>
  <dc:description xsi:type="xcriTerms:structure" />
  <dc:description xsi:type="xcriTerms:metadataKeywords">S100, Science: a
    foundation course, Undergraduate course, Open University</dc:description>
  <dc:description>Replacement type: Not a replacement.</dc:description>
  <dc:identifier>S100</dc:identifier>
  <dc:identifier>PLANET course version: 1.1</dc:identifier>
  <dc:subject />
```

```
<dc:title>Science: a foundation course</dc:title>
<mlo:url>http://data.open.ac.uk/page/course/s100</mlo:url>
<xcri:applicationProcedure>This course is closed and no longer in
  presentation.</xcri:applicationProcedure>
<mlo:assessment />
<xcri:learningOutcome />
-
= <mlo:qualification>
  <dc:identifier>OU Level 1 Introductory</dc:identifier>
  <dc:title>OU Level 1 Introductory</dc:title>
  <dcterms:educationLevel>Introductory</dcterms:educationLevel>
  <dcterms:type>OU module, Subtype: Standard module</dcterms:type>
  <xcri:awardedBy>The Open University</xcri:awardedBy>
  </mlo:qualification>
-
= <mlo:qualification>
  <dc:title />
  <dcterms:educationLevel>HEQF Level 4</dcterms:educationLevel>
  </mlo:qualification>
-
= <mlo:qualification>
  <xcri:abbr>BSc (Hons)</xcri:abbr>
  <dcterms:educationLevel>Undergraduate</dcterms:educationLevel>
  </mlo:qualification>
-
= <mlo:qualification>
  <dc:description>May lead to a variety of qualifications</dc:description>
  <dcterms:educationLevel>Undergraduate</dcterms:educationLevel>
  </mlo:qualification>
-
= <mlo:credit>
  <credit:scheme>ECTS</credit:scheme>
  <credit:level>1</credit:level>
  <credit:value>30</credit:value>
  </mlo:credit>
-
= <mlo:credit>
  <credit:scheme>The Open University</credit:scheme>
  <credit:level />
  <credit:value>60</credit:value>
  </mlo:credit>
-
= <xcri:presentation>
  <dc:date>1971</dc:date>
  <dc:description />
  <dc:identifier xsi:type="courseDataProgramme:internalID">S100
    1971</dc:identifier>
  <dc:title>Science: a foundation course</dc:title>
  <mlo:start dtf="1971-02-01">1971-02-01</mlo:start>
  <xcri:end dtf="1971-12-31">1971-12-31</xcri:end>
```

```
<mlo:duration />
<xcri:applyTo>http://www.open.ac.uk/library/library-resources/the-open-
  university-archive</xcri:applyTo>
<xcri:studyMode identifier="FL">Flexible</xcri:studyMode>
<xcri:attendanceMode identifier="DS">Distance without
  attendance</xcri:attendanceMode>
<xcri:attendancePattern identifier="CS">Customised</xcri:attendancePattern>
<mlo:languageOfInstruction type="iso639-1">en</mlo:languageOfInstruction>
<xcri:languageOfAssessment type="iso639-1">en</xcri:languageOfAssessment>
<mlo:places>unknown</mlo:places>
<mlo:cost>unknown</mlo:cost>
<xcri:age>18+</xcri:age>
  </xcri:presentation>
  -
  -
  .
  .
  .
  .
  .
= <xcri:presentation>
  <dc:date>1978</dc:date>
  <dc:description />
  <dc:identifier xsi:type="courseDataProgramme:internalID">S100
    1978</dc:identifier>
  <dc:title>Science: a foundation course</dc:title>
  <mlo:start dtf="1978-02-01">1978-02-01</mlo:start>
  <xcri:end dtf="1978-12-31">1978-12-31</xcri:end>
  <mlo:duration />
  <xcri:applyTo>http://www.open.ac.uk/library/library-resources/the-open-
    university-archive</xcri:applyTo>
  <xcri:studyMode identifier="FL">Flexible</xcri:studyMode>
  <xcri:attendanceMode identifier="DS">Distance without
    attendance</xcri:attendanceMode>
  <xcri:attendancePattern identifier="CS">Customised</xcri:attendancePattern>
  <mlo:languageOfInstruction type="iso639-1">en</mlo:languageOfInstruction>
  <xcri:languageOfAssessment type="iso639-1">en</xcri:languageOfAssessment>
  <mlo:places>unknown</mlo:places>
  <mlo:cost>unknown</mlo:cost>
  <xcri:age>18+</xcri:age>
    </xcri:presentation>
    </xcri:course>
    </xcri:provider>
    </xcri:catalog>
  </foxml:xmlContent>
  </foxml:datastreamVersion>
  </foxml:datastream>
  -
= <foxml:datastream ID="RELS-EXT" STATE="A" CONTROL_GROUP="X"
  VERSIONABLE="true">
```

```
-
=<foxml:datastreamVersion ID="RELS_EXT.0" LABEL="RELS-EXT" CREATED="2013-04-
22T10:16:29.631Z" MIMETYPE="text/xml" SIZE="42046">
-
=<foxml:xmlContent>
-
=<rdf:RDF xmlns:dc="http://purl.org/dc/elements/1.1/"
xmlns:fedora="info:fedora/fedora-system:def/relations-external#" xmlns:fedora-
model="info:fedora/fedora-system:def/model#"
xmlns:oai_dc="http://www.openarchives.org/OAI/2.0/oai_dc/"
xmlns:dcterms="http://purl.org/dc/terms/"
xmlns:ebucore="http://www.ebu.ch/metadata/ontologies/ebucore/ebucore#"
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
-
=<rdf:Description rdf:about="info:fedora/xcri:S100">
<dcterms:isReplacedBy rdf:resource="info:fedora/xcri:S101" />
<fedora:isMemberOf rdf:resource="info:fedora/model:module" />
<fedora:isMemberOfCollection rdf:resource="info:fedora/collect:cm" />
<fedora:hasPart rdf:resource="info:fedora/study:206652" />

.
.
.
.

<fedora:hasPart rdf:resource="info:fedora/study:419629" />
<fedora:hasPart rdf:resource="info:fedora/supp:8df997e8-665d-11e2-aa70-
eee94ec0690c" />

.
.
.

<fedora:hasPart rdf:resource="info:fedora/supp:8e033ec4-665d-11e2-aa70-
eee94ec0690c" />
<fedora:hasPart rdf:resource="info:fedora/video:JOUZ382E" />

.
.
.
.
.

<fedora:hasPart rdf:resource="info:fedora/audio:FM116J" />

.
.
.
```

```
<fedora:hasPart rdf:resource="info:fedora/S100_unit:Introduction" />
<fedora:hasPart rdf:resource="info:fedora/S100_unit:01" />
```

.

.

.

```
<fedora:hasPart rdf:resource="info:fedora/S100_unit:34" />
  </rdf:Description>
  </rdf:RDF>
  </foxml:xmlContent>
  </foxml:datastreamVersion>
  </foxml:datastream>
  </foxml:digitalObject>
```


8.5 Fedora Sample Record for Main Text

```
<foxml:digitalObject xmlns:foxml="info:fedora/fedora-system:def/foxml#"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" VERSION="1.1"
PID="study:206661" xsi:schemaLocation="info:fedora/fedora-system:def/foxml#
http://www.fedora.info/definitions/1/0/foxml1-1.xsd">
<foxml:objectProperties>
<foxml:property NAME="info:fedora/fedora-system:def/model#state"
VALUE="Active"/>
<foxml:property NAME="info:fedora/fedora-system:def/model#label"
VALUE="study:206661"/>
<foxml:property NAME="info:fedora/fedora-system:def/model#ownerId"
VALUE="xxx"/>
<foxml:property NAME="info:fedora/fedora-system:def/model#createdDate"
VALUE="2013-05-02T15:24:12.952Z"/>
<foxml:property NAME="info:fedora/fedora-system:def/view#lastModifiedDate"
VALUE="2013-05-02T15:24:12.952Z"/>
</foxml:objectProperties>
<foxml:datastream ID="AUDIT" STATE="A" CONTROL_GROUP="X"
VERSIONABLE="false">
<foxml:datastreamVersion ID="AUDIT.0" LABEL="Audit Trail for this object"
CREATED="2013-05-02T15:24:12.952Z" MIMETYPE="text/xml"
FORMAT_URI="info:fedora/fedora-system:format/xml.fedora.audit">
<foxml:xmlContent>
<audit:auditTrail xmlns:audit="info:fedora/fedora-system:def/audit#">
<audit:record ID="AUDREC1">
<audit:process type="Fedora API-M"/>
<audit:action>ingest</audit:action>
<audit:componentID/>
<audit:responsibility>xxx</audit:responsibility>
<audit:date>2013-05-02T15:24:12.952Z</audit:date>
<audit:justification>
Ingested from local file /study_206661.xml
</audit:justification>
</audit:record>
</audit:auditTrail>
</foxml:xmlContent>
</foxml:datastreamVersion>
</foxml:datastream>
<foxml:datastream ID="DC" STATE="A" CONTROL_GROUP="X" VERSIONABLE="true">
<foxml:datastreamVersion ID="DC1.0" LABEL="Dublin Core Record for this
object" CREATED="2013-05-02T15:24:12.970Z" MIMETYPE="text/xml"
FORMAT_URI="http://www.openarchives.org/OAI/2.0/oai_dc/" SIZE="485">
<foxml:xmlContent>
<oai_dc:dc xmlns:oai_dc="http://www.openarchives.org/OAI/2.0/oai_dc/"
xmlns:dc="http://purl.org/dc/elements/1.1/"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai_dc/
http://www.openarchives.org/OAI/2.0/oai_dc.xsd">
<dc:title/>
<dc:format/>
<dc:identifier>study:206661</dc:identifier>
<dc:identifier/>
<dc:identifier/>
<dc:source/>
</oai_dc:dc>
</foxml:xmlContent>
```

```
</foxml:datastreamVersion>
</foxml:datastream>
<foxml:datastream ID="MODS" STATE="A" CONTROL_GROUP="X" VERSIONABLE="true">
<foxml:datastreamVersion ID="MODS.0" LABEL="MODS" CREATED="2013-05-
02T15:24:12.971Z" MIMETYPE="text/xml" SIZE="2947">
<foxml:xmlContent>
<mods:mods xmlns:mods="http://www.loc.gov/mods/v3"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" version="3.4"
xsi:schemaLocation="http://www.loc.gov/mods/v3
http://www.loc.gov/standards/mods/v3/mods-3-4.xsd">
<mods:genre authority="local">instructional material:textbook</mods:genre>
<mods:titleInfo>
<mods:nonSort>The</mods:nonSort>
<mods:title>
periodic table and chemical bonding ; Ions in solution
</mods:title>
</mods:titleInfo>
<mods:titleInfo type="alternative">
<mods:title>Ions in solution</mods:title>
</mods:titleInfo>
<mods:titleInfo type="alternative">
<mods:title>Science foundation course</mods:title>
</mods:titleInfo>
<mods:name type="corporate">
<mods:namePart>Open University S100/Unit 8 and 9</mods:namePart>
</mods:name>
<mods:typeOfResource>text</mods:typeOfResource>
<mods:originInfo>
<mods:place>
<mods:placeTerm type="code" authority="marccountry">enk</mods:placeTerm>
</mods:place>
<mods:place>
<mods:placeTerm type="text">Milton Keynes</mods:placeTerm>
</mods:place>
<mods:publisher>Open University Press</mods:publisher>
<mods:dateIssued>1971</mods:dateIssued>
<mods:dateIssued encoding="marc">1971</mods:dateIssued>
<mods:issuance>monographic</mods:issuance>
</mods:originInfo>
<mods:language>
<mods:languageTerm authority="iso639-2b"
type="code">eng</mods:languageTerm>
</mods:language>
<mods:physicalDescription>
<mods:form authority="marcform">print</mods:form>
</mods:physicalDescription>
<mods:note>Units 8 & 9 of S100 Science foundation course</mods:note>
<mods:note type="version identification">LTO tape
number:LTO0160</mods:note>
<mods:note type="version identification">Hard drive
number:HDR0112</mods:note>
<mods:note>
Open University Module S100 Science foundation course
</mods:note>
<mods:subject authority="lcsh">
<mods:topic>Chemical elements</mods:topic>
</mods:subject>
<mods:subject authority="lcsh">
```

Project Identifier:
Version: 1.0
Contact: Sarah Brown
Date: 17/09/13

```
<mods:topic>Chemical bonds</mods:topic>
</mods:subject>
<mods:subject authority="lcsb">
<mods:topic>Ions</mods:topic>
</mods:subject>
<mods:classification authority="ddc">500</mods:classification>
<mods:classification authority="">S100</mods:classification>
<mods:relatedItem type="series">
<mods:titleInfo>
<mods:title>OU Study Materials</mods:title>
</mods:titleInfo>
</mods:relatedItem>
<mods:identifier type="isbn">0335020038</mods:identifier>
<mods:recordInfo>
<mods:recordContentSource
authority="marcorg">UkMkOU</mods:recordContentSource>
<mods:recordCreationDate encoding="marc">010504</mods:recordCreationDate>
<mods:recordChangeDate
encoding="iso8601">20120712094829.0</mods:recordChangeDate>
<mods:recordIdentifier>206661</mods:recordIdentifier>
<mods:recordOrigin>
Converted from MARCXML to MODS version 3.4 using MARC21slim2MODS3-4.xsl
(Revision 1.79 2012/07/11)
</mods:recordOrigin>
</mods:recordInfo>
</mods:mods>
</foxml:xmlContent>
</foxml:datastreamVersion>
</foxml:datastream>
<foxml:datastream ID="RELS-EXT" STATE="A" CONTROL_GROUP="X"
VERSIONABLE="true">
<foxml:datastreamVersion ID="RELS_EXT.0" LABEL="RELS-EXT" CREATED="2013-05-
02T15:24:12.971Z" MIMETYPE="text/xml" SIZE="5205">
<foxml:xmlContent>
<rdf:RDF xmlns:dc="http://purl.org/dc/elements/1.1/"
xmlns:fedora="info:fedora/fedora-system:def/relations-external#"
xmlns:fedora-model="info:fedora/fedora-system:def/model#"
xmlns:oai_dc="http://www.openarchives.org/OAI/2.0/oai_dc/"
xmlns:ebucore="http://www.ebu.ch/metadata/ontologies/ebucore/ebucore#"
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
<rdf:Description rdf:about="info:fedora/study:206661">
<fedora:isMemberOf rdf:resource="info:fedora/model:study"/>
<fedora:isMemberOfCollection rdf:resource="info:fedora/collect:cm"/>
<fedora:isPartOf rdf:resource="info:fedora/xcri:S100"/>
<fedora:hasPart
rdf:resource="info:fedora/page:63e2d7e9872ba9078bee90f86357449f342ee74b"/>
.
.
.
<fedora:hasPart
rdf:resource="info:fedora/page:bb944272d7dc8d3a2891e19095bc82d5a3e36960"/>
</rdf:Description>
</rdf:RDF>
</foxml:xmlContent>
</foxml:datastreamVersion>
</foxml:datastream>
</foxml:digitalObject>
```

8.6 Module profile

Desired label	Linked Data mapping	URIs	Schema Definition	
Module Code	<i>aiiso:module</i> aiiso:code	http://purl.org/vocab/aiiso/schema#code	The code used by an Institution to refer to a KnowledgeGrouping or Organization.	S100 Science: A Foundation Course
Course Type	<i>xcricourse</i> mlo:assessment	http://purl.org/net/mlo/assessment	The nature or genre of the resource.	4
Course SubType	<i>xcricourse</i> dc:type	http://purl.org/dc/terms/type	The nature or genre of the resource.	50 Standard module
Course version type	<i>xcricourse</i> dc:type	http://purl.org/dc/terms/type	The type of module the resource form	O Openings course
Course url	<i>xcricourse</i> mlo:url	http://purl.org/net/mlo/url	A link to a web resource that provides an alternate representation of the resource	http://data.open.ac.uk/page/course/m366.html
Admin				
Course Status	crsw:isTaught Present	http://courseware.rkbexplorer.com/ontologies/courseware#is-taught-present	NO DEFINITION IN RDF DOC	No longer current=False
Course Life	<i>xcricourse</i> dc:date	http://purl.org/dc/terms/date	Point or period of time associated with an event in the lifecycle of the resource. Date may be used to express temporal information at any level of granularity. Recommended best practice is to use an encoding scheme, such as the W3CDTF profile of ISO 8601 [W3CDTF]	8
Faculty	<i>aiiso:faculty</i>	http://purl.org/vocab/aiiso/schema#Faculty	Faculty: a group of people recognised by an organization as forming a cohesive group referred to by the organization as a faculty.	Faculty of Science
	<i>aiiso:part of</i> <i>fedora:isPartOf</i>	http://purl.org/vocab/aiiso/schema#part_of http://info.fedora/fedora-system:def/relations-external#isPartOf	partOf: is part of the Open University.	The Open University
	dc:contributor	http://purl.org/dc/elements/1.1/contributor	dc:contributor: Faculty is the contributor to the module	Faculty of Science

Subunit	aiiso:department	http://purl.org/vocab/aiiso/schema#Department	A Department is a group of people recognised by an organization as forming a cohesive group referred to by the organization as a department.	Department of Classical Studies
	aiiso:part of fedora:isPartOf	http://purl.org/vocab/aiiso/schema#part_of_info:fedora/fedora-system:def/relations-external#isPartOf	partOf: is part of the aiiso:faculty.	Faculty of the Arts
Programme	aiiso:programme	http://purl.org/vocab/aiiso/schema#Programme	A Programme is a KnowledgeGrouping that represents a cohesive collection of educational material referred to by the owning organization as a programme.	6 Arts and Humanities
	mlo:hasPart fedora:hasPart	http://purl.org/net/mlo/hasPart_info:fedora/fedora-system:def/relations-external#hasPart	This can be used to represent part-whole relations, such as the relationship between organisations and departments, or between programmes and constituent units. Content should be a URI conforming to IETF\RFC2396.	S100 Science: A Foundation Course
Version information				
This module has the following versions	dcterms:hasVersion	http://purl.org/dc/terms/hasVersion		None
Provider	xcrl:provider	http://xcrl.org/profiles/catalog/1.2/provider	Providers are organisations that offer one or more courses.	The Open University
	aiiso:organization	http://purl.org/vocab/aiiso/schema#organization		
	foaf:organization	http://xmlns.com/foaf/0.1/Organization		
	crsw:taught-at	http://www.aktors.org/ontology/portal#Organization		
Basic Data				
Title	xcrl:course dcterms:title	http://purl.org/dc/terms/title	A name given to the resource.	Science: A Foundation Course
Short Title	xcrl:course dcterms:alternative	http://xcrl.org/profiles/catalog/1.2/abbr	An abbreviated form of the title of the qualification	Science: A Foundation Course
Description	xcrl:course dcterms:description	http://purl.org/dc/terms/description	An account of the resource.	-

OU Level	<i>xcricourse</i> dcterms:educationalLevel	http://purl.org/dc/terms/educationalLevel	A class of entity, defined in terms of progression through an educational or training context, for which the described resource is intended.	1
Level 1 type	<i>xcricourse</i> dcterms:type	http://purl.org/dc/terms/type		Key introductory
Scottish Level	<i>xcricourse</i> dcterms:educationalLevel	http://purl.org/dc/terms/educationalLevel	A class of entity, defined in terms of progression through an educational or training context, for which the described resource is intended.	9
FHEQ Level	<i>xcricourse</i> dcterms:educationalLevel	http://purl.org/dc/terms/educationalLevel	A class of entity, defined in terms of progression through an educational or training context, for which the described resource is intended.	4
OU Credits	<i>xcricredit</i> <i>xcricreditscheme</i> <i>xcricreditlevel</i> <i>xcricreditvalue</i>	http://purl.org/net/mlo/credit http://purl.org/net/cm/scheme http://purl.org/net/cm/level http://purl.org/net/cm/value	The number of credits that have been obtained or can be obtained.	60
ECTS Credits	<i>xcricredit</i> <i>xcricreditscheme</i> <i>xcricreditlevel</i> <i>xcricreditvalue</i>	http://purl.org/net/mlo/credit http://purl.org/net/cm/terms/ECTS http://purl.org/net/cm/level http://purl.org/net/cm/value	The number of credits that have been obtained or can be obtained.	30
Assessed Module	<i>xcripresentation</i> mlo:assessment	http://purl.org/net/mlo/assessment	Assessment strategy: a description of the broad approach to assessment used in the Learning Opportunity.	
	crsw:has-assessment-method	http://courseware.rkbexplorer.com/ontologies/courseware#has-assessment-method		Yes
Taught Module	<i>xcripresentation</i> mlo:assessment	http://purl.org/net/mlo/assessment		

	crsw:has-assessment-method	http://courseware.rkbexplorer.com/ontologies/courseware#has-assessment-method		Yes
Study Hours	xcricourse dcterms:description type="studyHours"	http://xcric.org/profiles/catalog/terms		30
	crsw:hours-of-personal-study	http://courseware.rkbexplorer.com/ontologies/courseware#hours-of-personal-study		
Provided Resources	crsw:has-courseware	http://www.aktors.org/ontology/portal#Publication-Reference		
	dcterms:description type=provided Resource	http://purl.org/dc/terms/description http://xcric.org/profiles/catalog/terms		
Set book	xcricourse dcterms:description xsi:type="xcricTerms:indicative Resource"	http://purl.org/dc/terms/description	A set book identified for the module	
Curriculum				
Qualification Contribution	<i>xcricourse</i> dcterms:alternative	http://purl.org/dc/terms/alternative	Description may include but is not limited to: an abstract, a table of contents, a graphical representation, or a free-text account of the resource.	
Academic Subject Category	<i>xcricourse</i> dcterms:subject type="sotonJA"	http://purl.org/dc/terms/subject	The topic of the resource.	

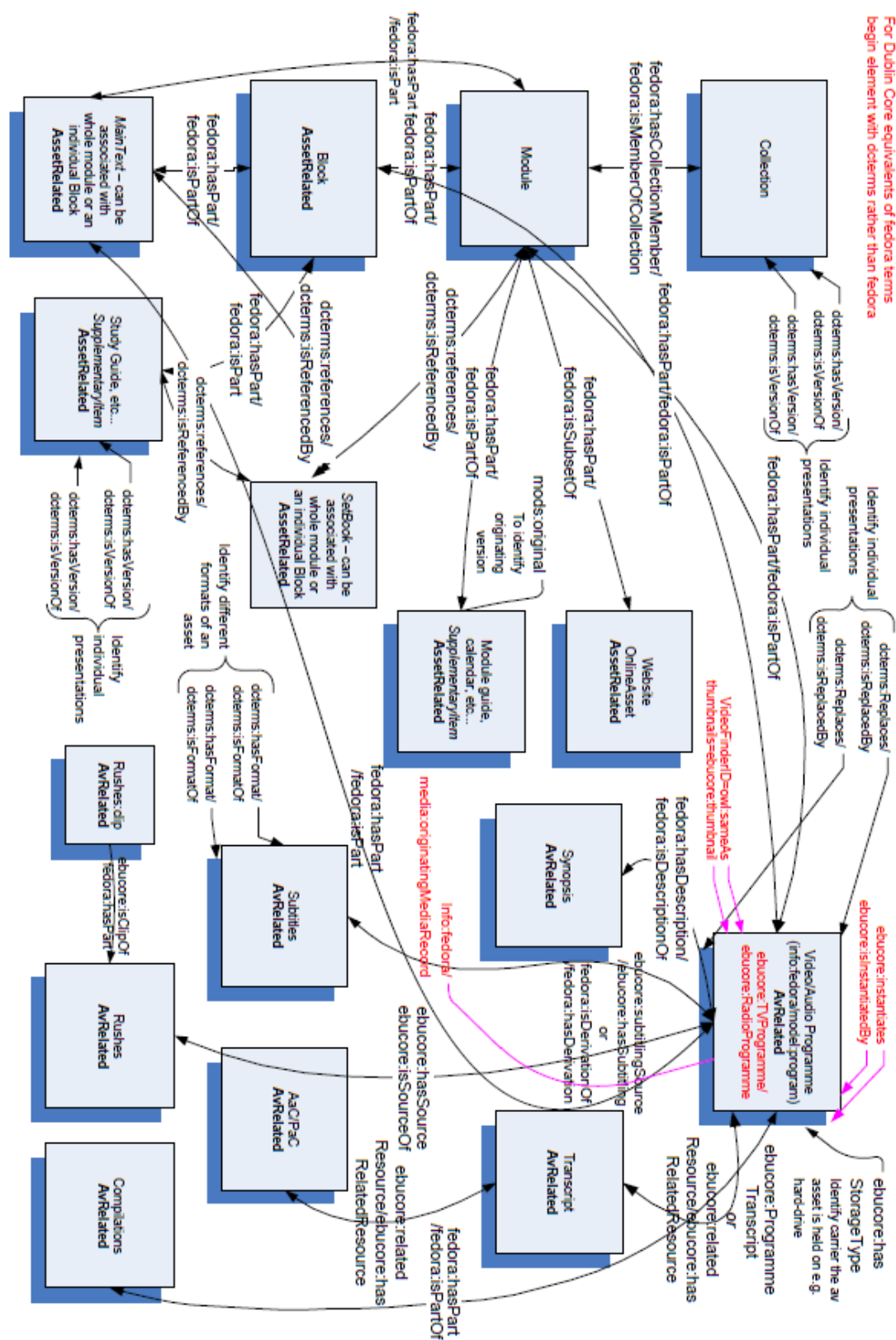
	CS"			
Replaced Module		http://purl.org/dc/terms/replaces	A related resource that is supplanted, displaced, or superseded by the described resource.	S102
Replacement Description	dcterms:description	http://purl.org/dc/terms/description	A related resource that is supplanted, displaced, or superseded by the described resource.	
Registration				
Attendance mode	<i>mlo:engagement</i> xcric.attendanceMode	http://xcric.org/profiles/catalog/1.2/attendanceMode	The type of location at which the student will undertake the learning opportunity, for example distance learning, campus-based, work-based, or online.	
Attendance Pattern	<i>mlo:engagement</i> xcric.attendancePattern	http://xcric.org/profiles/catalog/1.2/attendancePattern	The period in the day and/or frequency during which attendance at a venue is required (if any), for example evenings, daytime, weekends.	CS Customised
Study Mode	<i>xcric:engagement</i> xcric.studyMode	http://xcric.org/profiles/catalog/1.2/studyMode	A general expression of the overall amount of the student's time that is devoted to the learning opportunity, as defined by the provider.	PT Part time
Prior study requirements	<i>xcric.course</i> mlo.prerequisite	http://purl.org/net/mlo/prerequisite	A prerequisite or entry requirement for accessing a course or presentation. See [EN 15982].	
Geographical areas for registration	mlo.location type="ISO3166-1" type="M49"	http://purl.org/net/mlo/location	The spatial location of the Learning Opportunity Provider or Learning Opportunity Instance	
Presentations			Relations	
Presentation Year	<i>xcric.engagement</i> mlo.specifies	http://purl.org/net/mlo/specifies	IN XCRI-CAP THIS HAS BEEN Elided by the containment of Presentation elements in Course elements rather than by explicit association	1971, 1972, 1973, 1974, 1975, 1976, 1977, 1978
Start date	<i>xcric.presentation</i> mlo.start	http://purl.org/net/mlo/start	The date at which a presentation begins. See [EN 15982].	"30-Jan-1999"

Finish Date	<i>xc:presentation</i> xc:end	http://xc:org/profiles/catalog/1.2/end	The date or time at which the presentation of a course finishes.	"30-Sep-2007"
Study weeks	<i>xc:presentation</i> mlo:duration	http://purl.org/net/mlo/duration	A duration of a presentation	50
Module edition	dc:terms.hasVersion	http://purl.org/dc/terms/hasVersion		1
Qualification Links				
Qualification code	<i>xc:qualification</i> dc:identifier	http://purl.org/dc/terms/identifier	An unambiguous reference to the resource within a given context.	B16
Qualification Title	<i>xc:qualification</i> dc:title	http://purl.org/dc/terms/title		
Qualifications summary	<i>xc:qualification</i> dc:description	http://purl.org/dc/terms/description		
Subject Links				
OU Subject Category	<i>xc:course</i> dc:subject	http://purl.org/dc/terms/subject	The topic of the resource.	
OU Subject Name	<i>xc:course</i> dc:subject	http://purl.org/dc/terms/subject	The topic of the resource.	
Module Keywords	<i>xc:course</i> dc:subject	http://purl.org/dc/terms/subject	The topic of the resource.	
JACS Code	<i>xc:course</i> dc:subject type="JACS3"	http://purl.org/dc/terms/subject	The academic discipline of the resource	B140 Neuroscience
Voyager fields				
Topical Term	<i>xc:course</i> dc:subject type="LCSH"	http://purl.org/dc/terms/subject	The topic of the resource.	science
Language	<i>xc:presentation</i> mlo:languageOfInstruction type="ISO639-1"	http://purl.org/net/mlo/languageOfInstruction	A language in which the presentation is available to be taught.	eng
Course Lifespan	<i>xc:course</i> dc:date type="W3CDTF"	http://purl.org/dc/terms/date	Point or period of time associated with an event in the lifecycle of the resource. Date may be used to express temporal information at any level of granularity. Recommended best practice is to use an encoding scheme, such as the W3CDTF profile of ISO 8601 [W3CDTF]	1999-2008

Project Identifier:
Version: 1.0
Contact: Sarah Brown
Date: 17/09/13

Collection	dc.collection	http://purl.org/dc/terms/collection		
------------	---------------	---	--	--

8.7 OU study material object model



8.8 Draft Open University Digital Preservation Policy (as of 26/06/2013)

This document is currently being consulted on within The Open University.

Open University Digital Preservation Policy Revised Draft

1. Principle Statement
2. What is digital preservation and why is it important?
3. Context
4. Procedural Accountability and Governance
5. Scope
6. Preservation principles

Appendices:

- A. Related Policies, strategies and documents

Document History

<u>Date</u>	<u>Version</u>	<u>Description</u>	<u>Author / Modified</u>
July-August 2012	0.1	Structure and initial thoughts	Ruth Cammies /Liz Mallett
25.01.2013	0.6	Revised with suggestions from University Records Manager	Sam Mansfield /Ruth Cammies
21.02.2013	0.7	Revised after comments from Library Services staff	Liz Mallett
04.03.2013	1.0	Revised after comments from Library Leadership Team	Liz Mallett
11.06.2013	2.0	Revised after comments from stakeholders	Ruth Cammies/ Liz Mallett

Principle Statement

The Open University is a world leader in distance learning and has been a pioneer in developing methods for supported open learning since its creation in 1969. Library Services is at the heart of the Open University learning experience and has a mission to collect, actively manage, preserve and provide access to unique University assets (both physical and digital). The driving principle of the digital preservation policy is to ensure the on-going accessibility and usability of selected content created by the Open University.

What is Digital Preservation and why is it important?

Digital Preservation is defined as “A series of managed activities necessary to ensure continued access to digital materials for as long as necessary. It refers to all of the actions required to maintain access to digital materials beyond the limits of media failure or technological change...”

(Beagrie and Jones 2002)

Digital preservation brings a good *return on investment* for the Open University for the following reasons:

- It underpins the ability to manage business risk in a digital environment – helping to deliver both business continuity & digital continuity
- It helps to ensure legal compliance
- It underpins the ability to deliver responsive services to digital scholars
- It ensures the record of the university's teaching methods and research are not lost

Digital preservation requires active, ongoing monitoring and undertaking activities to ensure digital objects continue to be authentic, understandable and accessible.

What DP is *not*:

It is important to note that, although an important component of a DP strategy, backing up data and use of checksums alone are not long term maintenance techniques for good preservation. They are indeed a minimum requirement for disaster recovery or business continuity, but they do not preserve the meaning associated with the values in the datasets and the ability to understand, read or re-use the digital object in the future.

A more detailed description of digital preservation and the benefits of undertaking digital preservation actions can be found in the *Digital Preservation White Paper*⁵⁹

Context

A number of related policies and strategies should be considered alongside the Digital Preservation Policy. They are listed in appendix A.

Procedural Accountability and Governance

- The Director of Library Services is the strategic accountable officer for University Archives, of which Digital Preservation is a part.
- The University Archivist is responsible for the stewardship of archive materials and associated services within the University, including digital preservation.
- Decisions regarding selection and preservation of major types of materials, will be based on wide consultation with Units and will take into account resource requirements, legal implications and best practice within HE.

⁵⁹ *Digital Preservation White Paper* (2011, Lorna Maguire, former Head of Information Management, Open University Library Services).

- The final decision regarding selection of materials for digital preservation will rest with...[to be confirmed].

Scope

This policy and related guidance covers:

- A selection of “born digital” content created by The Open University which is deemed worthy of long term preservation according to the selection criteria.
- A selection of digital surrogate files created (i.e. digitised) from “original” (non-digital) objects/documents, e.g. transcripts, photographs, video, audio, etc.

Materials will be included in accordance with the University Records Retention Schedule and Institutional Archives Policy (currently in draft form).

A separate strategy document regarding selection criteria – and the operational strategies for preservation – is being developed by Library Services, in consultation with CAUs and other administrative Units.

Library Services seeks to make preserved materials and associated metadata as accessible as possible – but does not seek to duplicate or divert usage from external access platforms such as iTunesU, OpenLearn or YouTube or act contrary to the licensing/sale work of the Business Development Unit. While archive copies of objects may be held within a preservation repository, access will be directed to the appropriate external platform where appropriate.

Out of Scope

The vast majority of digital files created by The Open University as part of its day to day business will not be preserved long-term. The Information Management Services intranet site contains guidance for units on managing non-permanent electronic files.

Preservation Principles

Select the right materials

Selection of digital content for preservation is key to ensuring efficiency and return on investment – not all digital content could be or *should be* preserved long-term.

Library Services staff will work with colleagues from other units to ensure that appropriate material is selected and preserved.

Obtain a copy of materials for preservation at an appropriate time

Digital Preservation actions need to “start” at an early stage of an object’s life. Waiting until an object is no longer current (or in active use) before beginning preservation actions increases the risk that the object is no longer accessible. Essential information about the object may also have been lost.

Library Services will work with other Units to ensure that necessary actions for preservation are taken at the appropriate time in the “life” of digital content. Wherever possible, operational workflows will be created to ensure materials are regularly captured to minimise additional work.

Capture the context/provenance of materials

In order for digital materials to be effectively preserved and re-used, sufficient information or “metadata” needs to be recorded. Such data captures key information regarding the material – such as date of creation, author, owner etc. Metadata standards exist which aim to make such data consistent with other collections of material and to ensure that the right data is being captured and recorded in the right way.

Without metadata the objects are at risk of the following:

- Not being findable by search engines
- Users being unsure of the version of the object, where it came from, who created it, the software with which it was created and who can give permission to reuse it– leading to the object being unusable.

Ensure the authenticity/integrity of materials

Preserved materials must be actively monitored while being stored to check for errors. Digital content degrades over time. Regular integrity checks on digital objects – and subsequent actions should errors be discovered, ensure that materials remain authentic and unchanged from the original.

Manage the continued accessibility of digital content

To ensure materials remain accessible over time, strategies need to be in place to monitor the risk of format obsolescence or objects becoming inaccessible due to the use of proprietary formats and systems.

Appendix A

Related Policies, strategies and documents

University Archives Policy (revised draft)
Information and Records Management Policy
Open University Records Retention Schedule
Institutional Archives Policy (draft)
University Data Protection site
University Freedom of Information site

8.9 Draft Open University Digital Preservation Strategy and Selection Criteria (as of 26/06/2013)

This document is currently being consulted on within The Open University.

Open University Digital Preservation Strategy and Selection Criteria

Draft

1. Context
2. Digital Preservation Strategies
3. General selection criteria for digital archive content
4. Selection policies for types of Open University content
 - 4.1 Study materials:
 - 4.1.1 VLE Web sites
 - 4.1.2 Video and audio
 - 4.1.3 Text materials
 - 4.2 Research content:
 - 4.2.1 Research outputs (text-based)
 - 4.2.2 Research data
 - 4.3 Institutional Archive – business records
 - 4.3.1 Web content
 - 4.3.2 Communications (such as podcasts)
 - 4.3.3 Business records
 - 4.4 Informal learning content

Document History

<u>Date</u>	<u>Version</u>	<u>Description</u>	<u>Author / Modified</u>
20.06.2013	0.1	Initial draft	Ruth Cammies /Liz Mallett

1. Context

This document accompanies the Open University Digital Preservation Policy. The aim of the document is to describe the strategies which will be used to implement the policy and to provide selection criteria for digital archives for the Open University.

2. Digital Preservation Strategies

The Digital Preservation Policy will be implemented through the following strategies:

2.1 OUDL (Open University Digital Library)

Library Services' main strategy for digital preservation is to create a repository system with preservation functionality, which also provides access to materials.

The Open University Digital Library (OUDL) Project is creating a Fedora Commons system. The system will have the following main functions:

- Ability to store and manage digital archive materials of any format (dependant on available storage space)
- Ability to automatically capture technical metadata for preservation and re-use purposes
- Ability to perform integrity checks on items and report errors on digital content
- Completely adaptable metadata control to allow customised metadata profiles for different types of digital object according to standards
- Secure, managed ingest of materials
- Ability to manage and hold metadata for non-digital content outside the system
- Ability to link to other systems where necessary
- Sophisticated, faceted search functionality
- Ability to maintain complex semantic relationships between digital objects and surface these as linked data.
- Ability to preview items – including full archive audio/visual content
- Ability to enable different levels of access to different groups of people – including groups external to the University

The University Archive team will answer enquiries from staff and external researchers regarding OUDL and continue to add materials, update metadata and monitor materials for preservation purposes.

2.2 Monitor materials held in other University repositories/systems

Where there are items worthy of long-term/permanent preservation held in other University repositories/systems, the repository/system needs to be assessed regarding its ability to meet digital preservation standards and perform preservation tasks. Where repositories are assessed as not suitable for preservation, there should be consultation regarding whether the materials should be moved to the OUDL preservation repository.

2.3 Monitor the use of appropriate formats for preservation

University Archive staff will seek to monitor formats being used for objects selected for preservation and will migrate materials when necessary to more appropriate formats. Where possible, copies of materials in original formats will also be preserved alongside the new versions.

2.4 Digitisation

Digitisation is used as a preservation strategy when the original format of the materials is at risk due to format obsolescence or fragility. Digitisation is therefore used as a preservation strategy for analogue audio/visual materials where older formats become inaccessible. The University Archive

within Library Services has an annual programme of digitisation of audio/visual materials in obsolete formats.

Digitisation is also used to increase access to content – such as printed materials and photographs. Library Services activities include digitisation of non-obsolete materials such as printed study materials texts to increase access to materials. Digitised access copies will be preserved to ensure ongoing accessibility.

2.5 Advice and support for projects

When planning projects, the sustainability of project data and outputs beyond the project's life should be considered from the early stages. With research projects in particular, long term access and preservation arrangements for data and outputs may be a requirement of the funder.

The Research Data Management Project being run jointly by the Research School and Library Services is currently assessing the most appropriate workflows/systems for research data within the University. This strategy will be updated to reflect the outcomes of the project with regards to the digital preservation of research data.

2.6 Membership of Specialist Preservation Organisations and Standards Compliance

The University is a Board Member of the Digital Preservation Coalition⁶⁰ and considers it essential to adhere to international digital preservation standards upheld by the DPC and other bodies such as The National Archives⁶¹, The National Preservation Advisory Centre at the British Library⁶².

The University is also a member of PrestoCentre – 'a membership-driven organisation that brings together a global community of stakeholders in audiovisual digitisation and digital preservation to share, work and learn'.⁶³

In addition, the University is working with the Digital Curation Centre regarding research data management.

The OUDL project team are using the following standards in the design and development of the Fedora digital library system:

- OAIS⁶⁴ (Reference Model for an Open Archival Information System)
- TRAC⁶⁵ (Trustworthy Repositories Audit and Certification: Criteria and Checklist)

The OUDL system will use the following international metadata standards:

Content Metadata:

- [OAI-Dublin Core](#) and [Dublin Core qualified](#) terms
- [PB Core](#) and [EBU Core](#) (Audio-visual content)
- [MARC21](#) and [MODS](#) (Bibliographic texts and supplementary texts)
- [VRA](#) and [MIX](#) (still images)
- [EAD](#) (archive special collections)

Structural Metadata:

- [FOXML](#)
- [METS](#)

⁶⁰ www.dpconline.org

⁶¹ www.nationalarchives.gov.uk

⁶² www.bl.uk/blpac

⁶³ <http://www.prestocentre.org/about-us>

⁶⁴ <http://public.ccsds.org/publications/archive/650x0m2.pdf>

⁶⁵ http://www.crl.edu/sites/default/files/attachments/pages/trac_0.pdf

- [EAD](#)

Digital Preservation:

- [PREMIS](#)

2.7 Auditing and Risk Assessment

It is essential to routinely monitor digital preservation policies and practice to ensure that they remain up-to-date and address all requirements.

During the initial OUDL project the DRAMBORA interactive tool⁶⁶ will be used alongside the TRAC Trusted Repositories Assessment Criteria⁶⁷ to assess risks and create a collection of required documentation for effective digital preservation. The risk register and associated policies and procedures will be regularly reviewed and updated as necessary.

3. General selection criteria for digital archive content

Selection criteria have been adapted from the University Archive Policy. Content types are in order of priority, with the first type being the highest priority for selection.

- Priority type 1: (core business):
 - Archive materials which represent OU heritage
 - OU study materials and associated materials
 - Materials which reflect distance/e-learning pedagogy, specific teaching techniques
 - OU business/institutional archive
- Priority type 2: (Internal Research)
Selection, management and preservation of internally generated research data will be determined by the outcomes of the Research Data Management Project.
- Priority type 3: (External research donations)
Unique research archive material which has been offered to the University from external bodies or individual researchers. Selection will be based on a case by case basis, guided by defined criteria and direct relevance to approved areas of research and teaching.
- Priority type 4: ("other" offered donations)
Miscellaneous material which has been offered, which does not readily fall into types 1-3. This may be a collection of published/unpublished material, collected by an individual/researcher over time.

3.1 Assessment

Collections of priority types 2-4 will be assessed with the following criteria:

- Level of stakeholder endorsement - by the relevant Dean (or more than one Dean where necessary with multidisciplinary collections)
- Level of significance of the archive collection in matching research interests.
- Flexibility of the deposit agreement and the ability to (re) use content
- Level of accessibility to researchers/public – immediate/closed or restricted access
- Assessment of exceptional cases; reputation/political risk - such as bequests
- Level of ability to support and deliver digital scholarship needs
- Clarity of provenance/ownership/structure/context

⁶⁶ <http://www.repositoryaudit.eu/>

⁶⁷ http://www.crl.edu/sites/default/files/attachments/pages/trac_0.pdf

- Uniqueness – what level of duplication from elsewhere; photocopies /articles /website /published reports

Assessment of material for inclusion within the University Archive will also include considerations around costs, balanced against potential research value.

Each collection will be initially assessed on the following aspects:

- Cost of acquisition - activities associated with obtaining the material
- Cost of accession - activities associated with making the material accessible
- Cost of long term sustainability/preservation
- Service support

Data will also be collected on an on-going basis regarding:

- Levels of use
- Relevance to core business
- Cost of continued resource investment

4.Selection policies for types of Open University content

This section includes selection policies for some of the main types of digital materials from priority type 1 (core business) that will be preserved as part of the University Archive. The list is not exhaustive and will be expanded as required.

The selection policies are subject to regular review.

4.1 Study materials:

4.1.1 VLE Web sites and attached content

4.1.1a Existing archive collections:

Web archiving of VLE web sites within the University Archive started in 2012. Selected versions of available Moodle VLE websites are being copied from 2006 onwards.

4.1.1b Selection criteria for digital preservation:

The first, last and a mid-point presentation⁶⁸ of each module within the VLE will be captured using web harvesting methods and preserved and made accessible to staff within the OU Digital Library, using Wayback Machine software developed by the Internet Archive.

Sites will usually be harvested from the learn-acct site rather than the live VLE, to avoid any usage impact on the live site.

Preserved versions of VLE sites will not ordinarily include any discussion boards or forums in accordance with Data Protection Legislation.

The University Archive team is working with LTS to incorporate VLE archiving into the deletion workflow of sites from the VLE platforms.

Faculty staff are encouraged to contact the University Archive if versions of a module (other than the first, last and a mid point) are of particular historical significance or exemplify particular teaching methods, or if particular forum content should be captured and anonymised.

4.1.1c Method of capture:

⁶⁸ This is the current interim policy. Alternative suggestions are welcome.

The University Archive is using web harvesting Wget software to harvest materials and a local instance of the WayBack Machine to view the archived content. While this method effectively copies the structure and the look and feel of a site, it is possible that some content (such as content located on non-VLE sites or some types of dynamic content) may not be captured. As the archiving process for these sites has commenced in 2012, early Moodle 1 websites may appear differently to how they were originally presented to students, due to updates to the global style sheets for the VLE.

4.1.1d Details of digital content:

Materials will be stored in .warc format.

4.1.1e Storage, Preservation and Access

Preserved VLE sites will be:

- catalogued using relevant standards
- stored within the OUDL system
- searchable and available to staff using Wayback Machine software, through the OUDL system.

Modules on the VLE Moodle site will be deleted according to the agreed University retention period.

4.1.2 Video and Audio

4.1.2a Existing archive collections:

The University Archive maintains a comprehensive collection of video and audio materials dating from the first OU courses in 1971 onwards. There is currently an on-going annual programme to digitise materials on at-risk and inaccessible formats, for purposes of re-use and to preserve the unique teaching contribution of the Open University.

4.1.2b Selection Criteria for digital preservation:

For long-term preservation purposes, the University Archive seeks to preserve high-quality versions of completed content used within module materials, together with associated metadata to aid preservation, capture the full context of the asset and enable re-use.

There is recognition that non-final versions of assets may be of particular re-use value for a period of time. While the primary location of these assets should be within production systems, there may be an occasional requirement to include them within a preservation system for a period of time or permanently if judged to be of particularly high value.

4.1.2c Method of capture:

New video and audio content for study materials are currently transferred to the University Archive by OMU on hard drive. Metadata is captured by OMU staff and entered into the Videofinder system. Further work is required to review and formalise this process and to establish workflow processes between the OUDL system and production systems.

Any video or audio content currently sent to students on media such as DVD or CD are also transferred to the University Archive (via the University warehouse) and added to the Library Voyager catalogue. The cataloguing process for these items will also be reviewed to meet requirements of the OUDL service.

Legacy archive material is being digitised on an annual basis by Library Services. A digitise-on-demand service is also being offered where materials are not currently in digital format.

4.1.2d Details of digital content:

- Archive materials need to be stored in a format suitable for preservation. The University Archive is currently using .avi as a video preservation format, but this will be reviewed in consultation with colleagues across the university and with advice from external expert bodies.
- Access copies of final content will be stored in a low-res format (currently flash format is being used) – this will be reviewed regularly

4.1.2e Storage, Preservation and Access:

- Initially, large high resolution files will be stored off-line on hard drive and backed-up on LTO data tape. This practice will be reviewed regularly and the cost of large capacity server storage will be monitored. All offline storage will be regularly checked for errors and continued accessibility.
- Storage of preservation content (unless stored offline) will be based within the OUDL system. Where files are stored offline, preservation of the storage media will be managed within the OUDL system.
- Digital A/V archives will be searchable within OUDL and made available to staff. When appropriate they may be made available more widely.

4.1.3 Text Materials

4.1.3a Existing archive collections:

The University Archive maintains a comprehensive physical (non-digital) collection of printed study materials available to view by appointment. The Archive continues to collect full student packs of first and last presentations of modules, and any new materials produced in intervening years.

4.1.3b Selection Criteria for digital preservation:

The University Archive seeks to preserve digital text materials presented to students. Full digital copies of module texts will be retained from the first and last presentation, and any new materials (or new versions of materials) from intervening years.

4.1.3c Method of capture:

- Digitisation of physical material (to increase access to materials)
- Locating and pointing to existing digital copies of texts in university systems
- Harvesting digital texts from VLE sites

4.1.3d Details of digital content:

The University Archive will aim to preserve a copy of original formats of text materials where appropriate. Text materials will also be converted to high quality .tiff image files and .pdf/a with optical character recognition (OCR).

4.1.3e Storage, Preservation and Access

Preserved text materials will:

- be catalogued according to relevant standards
- be stored within the OUDL system where appropriate
- be searchable and made available to staff and may be made available more widely as appropriate.

4.2 Research content:

4.2.1 Research articles and papers:

Peer reviewed research outputs are currently listed or added to the University's ORO (Open Research Online) Eprints repository⁶⁹. The Open University Archive will work with the ORO team to establish preservation actions for materials held within the repository.

4.2.2 Research data:

⁶⁹ <http://oro.open.ac.uk/>

The Research Data Management Project is currently underway to consider the long-term strategy of data management, access and preservation for the University.

While this project is underway new offers of digital (or non-digital) research data collections will be considered with advice/support from the Dean of the relevant Faculty and considerations regarding cost, collection size and the nature of the material.

4.3 Institutional Archive – business records

4.3.1 Web content

4.3.1a Existing Archive Collections

The University Archive started to collect non-VLE websites in 2013.

4.3.1b Selection Criteria

The University Archive seeks to harvest web sites of major projects and University initiatives, and sites reflecting changes and development of the University on a case-by-case basis.

Web sites containing personal information will usually not be archived – unless there is a significant justification. In such cases all personal information will be anonymised/removed from the archive copy.

The University Archive welcomes requests/suggestions from staff for web sites that could be suitable for archiving and aims to actively select and capture suitable sites.

4.3.1c Method of capture:

Web sites will be captured using web harvesting software. An arrangement has been made with IT to log any live sites due to be harvested.

4.3.1d Details of digital content:

Content will be stored in .warc file format.

4.3.1e Storage, Preservation and Access:

Web sites will be preserved and searchable within the OUDL system and made available to staff using Wayback Machine software. Web sites may be made available more widely as appropriate.

4.3.2 University Communications and internal events recordings

4.3.2a Existing Archive Collections

The University Archive contains various series of physical (non-digital) printed reports and news-letters/magazines such as Open House, Sesame, printed press releases and annual reports.

The University Archive holds a tape-based collection of AV Department recordings.

4.3.2b Selection Criteria for digital preservation

As more regular communications are e-only, the University Archive aims to collect this type of material in digital form, to be preserved using OUDL. This could include text-based resources and also video/audio communications (such as podcasts) or web-based information.

The University Archive team will actively investigate suitable content to be collected, and also welcomes requests/suggestions from staff for digital reports/news-letters/magazines which could be suitable for archiving.

4.3.2c Method of capture

Procedures regarding the capture and storage of communications and internal recordings and related metadata will be written in consultation with colleagues across the university.

4.3.2d Details of digital content

Internal communications and events recordings may be in a variety of formats.

4.3.2e Storage, preservation and access

To be confirmed through consultation with colleagues.

4.3.3 Business Records

The University Archivist and University Records Manager are currently working to create a policy/procedure regarding digital institutional archives.

4.3.3a Existing Archive Collections

The University Archive holds a selection of physical (non-digital) business records, such as committee papers. A programme of work has also been undertaken to digitise selected high-level committee papers, such as Council and Senate.

4.3.3b Selection Criteria

Selection criteria for digital institutional archives will follow the Records Retention Schedule and the draft Institutional Archive Policy. An extract from the draft policy below gives a high-level guideline to materials that will and will not normally be considered for permanent/long-term preservation.

Materials will be considered as suitable for deposit within the Institutional archive if they meet one or more of the following criteria:

- Demonstrating major changes in the OU
- Showing evidence of major projects/functions/activities at a senior University level
- Showing a high level decision making process
- Containing material that reflect the views and activities of individuals who have placed a significant role in the development of the OU
- Reflecting changes in attitudes, or policies over time
- Promoting an understanding of issues unique to the OU
- Possessing aesthetic qualities that may be useful for exhibition or promotion of the OU.
- Are products of the University's teaching and learning function, i.e. learning materials, and significant records of their design.

Materials with the following characteristics are unlikely to be accepted:

- Operational papers (emails/meeting minutes etc) for routine activities of the University
- Collections with high levels of duplication
- Collections with a considerable number of documents created external to the University

4.3.3c Method of Capture

Procedures regarding the capture and storage of digital institutional archives and related metadata will be written in consultation with colleagues across the university.

4.3.3d Details of digital content

Digital institutional archives could be of any format.

4.3.3e Storage, Preservation and Access

To be confirmed.

4.4 Informal learning content

Including (but not limited to):

- OU/BBC (and other broadcasters) co-produced television/radio content
- Legacy OU non-teaching television/radio content (such as *Open Forum*)
- Other OU/third-party collaborative content
- Content commissioned for iTunesU, YouTube, other sites

4.4a Existing Archive Collections:

The University Archive holds a collection of non-teaching audio and video content dating from 1971 onwards, including such series as *Open Forum* and events like the first graduation ceremony. The collection is partially digitised. Additional items are digitised on-demand or as part of the annual digitisation work.

The University Archive holds tape copies of OU/BBC content such as *Coast*, *Child of Our Time* etc.

4.4b Selection Criteria:

The University Archive aims to preserve all informal learning materials presented to external audiences.

4.4c Method of Capture:

The University Archive seeks to work with the OMU and LTS to improve/formalise preservation workflows for digital non-teaching A/V materials – including capturing relevant metadata.

4.4d Details of Digital Content:

A copy of assets will be retained in the original (transfer) format where possible.

Video assets may also be transferred to uncompressed .avi format for preservation (to be reviewed).

Audio assets may be transferred to .wav for preservation

4.4e Storage, Preservation and Access:

- Records of assets will be stored within OUDL (large video assets may be stored offline on hard drive and LTO data tape or on the AV storage system when costs are less prohibitive – records will still be managed within OUDL)
- Where appropriate, contents will be made available to staff through OUDL – if/when the items are unavailable from the original access platform (such as Stadium/iTunesU). Where appropriate the OUDL system will link to the preferred access platform – such as iTunesU/YouTube. In these cases a preservation copy will still be stored but will not be accessible to staff or a wider audience.